uc3m | Universidad Carlos III de Madrid

Ph.D. Thesis

Essays on mixed-frequency and causality analysis in the frequency domain

Cleiton Guollo Taufemback

Advisor
Carlos Velasco Gómez, Ph.D.

Economics Department

Getafe, July/2018

Contents

List of Figures

List of Tables

Introduction

In this thesis, we present three different studies about mixed-frequency datasets and/or causality.

The literature of mixed-frequency econometrics is scarce of how discarding intermediate data, also known as temporal aggregation, impacts on estimation consistency. Thereby, in Chapter 1, we focus on showing how temporal aggregation leads to inconsistent least squares estimators whenever the DGP exhibits frequency dependent coefficients, FDC. The spectrum of a subsampled variable is equal to its folded original spectrum. As result, the low-frequency variable may present, for a given frequency, a mix of distinct linear relations. Furthermore, based on band spectrum regression, we propose a new method to disentangle the frequencies superposition and to circumvent the inconsistency problem. Under this methodology, it is also possible to test whether the coefficients are frequency dependent. We analyze stationary and nonstationary linear semiparametric/nonparametric models, as well as stationary efficiency. Our Monte Carlo simulations illustrate good finite sample properties in detecting the correct number of FDC. Finally, an empirical application of our method to quarterly GDP and US monthly indicators rejects the presence of a unique coefficient for all frequencies.

In Chapter 2, we propose a novel nonparametric frequency Granger-causality test. Before testing for causality absence of one series on to another, we apply a filtering process, removing any presence of reverse causality. Then, performing a local kernel regression for each frequency, we can test the hypothesis of noncausality from the distance between each estimate to zero. We provide asymptotic results for strict stationary series respecting α-mixing conditions. Monte Carlo experiments illustrate that our method has good finite sample properties, with overall comparable performance with other alternative methods present in the literature, and superior performance whenever the tested model presents smooth transition coefficients in the frequency domain. Finally, we test causality of term spread and money stock (M2) on real economic growth, as well as, between Monetary Policy Variables and Stock Prices.

In Chapter 3, we propose two novel nonparametric causality tests for mixed-frequency datasets. One based on least squares, LS, estimation and other based on the Hannan-Inefficient, HI, estimator. In our framework, the dependent variable is fitted by an increasing, with the sample size, number of leads and lags of the exogenous variable. The LS approach presents better results for testing causality at individual leads/lags, and the HI approach is superior for addressing the joint null hypothesis of non-causality from one series to another. Furthermore, we assume that the low-frequency variable is generated at high frequency, but some observations are systematically unobserved, rather than assume distinct generation frequencies as in MIDAS and mixed-frequency VAR models. Thus, our approach results in a more straightforward interpretation of causality. Finite sample simulations show good results in size and power for our causality tests and consistent estimation of leads and lags coefficients. Finally, we provide some empirical results on testing for no causality between GDP and US monthly indicators.

Chapter 1

Asymptotic behavior of temporal aggregation in mixed-frequency datasets

Abstract

The literature of mixed-frequency econometrics is scarce of how discarding intermediate data, also known as temporal aggregation, impacts on estimation consistency. Thereby, in this paper, we focus on showing how temporal aggregation leads to inconsistent least squares estimators whenever the DGP exhibits frequency dependent coefficients, FDC. The spectrum of a subsampled variable is equal to its folded original spectrum. As result, the low-frequency variable may present, for a given frequency, a mix of distinct linear relations. Furthermore, based on band spectrum regression, we propose a new method to disentangle the frequencies superposition and to circumvent the inconsistency problem. Under this methodology, it is also possible to test whether the coefficients are frequency dependent. We analyze stationary and nonstationary linear semiparametric/nonparametric models, as well as stationary efficiency. Our Monte Carlo simulations illustrate good finite sample properties in detecting the correct number of FDC. Finally, an empirical application of our method to quarterly GDP and US monthly indicators rejects the presence of a unique coefficient for all frequencies.

Keywords: Mixed-Frequency Data, Frequency Domain, Linear Models, Aliasing.
JEL classification: C22, C52.

1 Introduction

In a profound revision of mixed-frequency data, MFD, techniques, Foroni and Marcellino (2013) affirm that the predominant approach to deal with MFD is to discard intermediate data, also called temporal aggregation, despite the advance in the literature, e.g., Mariano and Murasawa (2010); Angelini et al. (2006); Bańbura and Modugno (2014); Ghysels et al. (2004a, 2007a); Bai et al. (2013a). Here, we understand as temporal aggregation both systematic sampling and non-overlapping aggregation. The first procedure skips stock observations of x_t that do not match in time with available observations of y_t and the second performs a continuum summation of a flow variable up to the release date. Please see Granger and Siklos (1995) and Hassler (2011) for extended discussion about temporal aggregation definition.

Moreover, Foroni and Marcellino (2013) argue that temporal aggregation has a potential loss of information since empirical papers have shown that MFD techniques improve forecast results. Based on a nonlinear model, Andreou et al. (2010) gave a theoretical indicative of temporal aggregation procedure problems. They showed that omitting the nonlinear term the least squares estimated coefficients might be inconsistent.

In this paper, using semiparametric and nonparametric linear models, we aim to establish that the subsampling process that occurs with the dependent variable, y_t, may affect the consistency of least squares linear estimators. The critical point of our analysis is to show how aliasing impacts when the model allows for frequency dependent coefficients, FDC. Notice that, for a single coefficient for all frequencies, the loss of information only affects the estimator efficiency, not its consistency.

Assuming that the DGP has coefficients that vary across frequencies, the two main results of this paper are: (1) regressing low-frequency series on a downsampled exogenous series result in inconsistent estimators or in the inability to recover all coefficients; (2) a new proposed method, based on band spectrum regression, can consistently recover distinct frequency coefficients.

Despite the literature about FDC started with Hannan (1963a) and Engle (1974), no paper, so far, has exploited any frequency domain approach as a possible solution for MFD. Considering that we are dealing with a sampling mismatch, frequency domain analyses together with Nyquist-Shannon Sampling theorem, see Shannon (1949), seems to be an intuitive field for new insights. For example in a quarterly-monthly scheme, one can interpret the quarterly series as a sub-sampled version of a possible non-observable monthly series. Although being impossible to recover missing values, the normalized spectrum of quarterly series is equal to the folded spectrum of the original monthly series, see Cooley et al. (1969). Understanding this relationship between low/high-frequency spectrum allows us to recover all coefficients consistently. As the spectrum of exogenous variables is unfolded, we can split it into frequency bands and then mimic the folding process that happened with the dependent variable. This process allows us to regress the aliased spectrum of the low-frequency variable over distinct frequency sets of the high-frequency variable periodogram.

It is important to mention that we assume that the same kind of variable, stock or flow, is present on both sides of the equal sign. Tiao (1972) presents some results of temporal aggregation of flow variables in the time domain. Hassler (2013) and Hassler and Tsai (2013) show the effects of subsampling in the frequency domain. According to their results, the spectrum of a flow variable can be written as a summation of the spectrum of its innovations. However, we treat the spectrum of a flow variable as it was a stock one. This loss of generality simplifies our notation and do not alter our findings. In Section 7, we discuss how to proceed to estimate linear models using stock and flow variables.

Some notations used through the paper: s means the ratio between the number of observations of high frequency, n, and low frequency variables, n_s; $z_{s,t}$, $n_s \times p$, represents the low frequency version of a high frequency variable z_t, $n \times p$. The variable Z_n represents the vector $\{z_t\}_{t=1}^n$, and $Z_{s,n}$ represents the vector $\{z_t\}$ sampled at $t = s, 2s, \ldots, n$. $w_z = WZ_n$, where W, $n \times n$, represents the discrete Fourier transform with row j given by $W_j = n^{-1/2}(1, e^{-i\lambda_j}, e^{-i2\lambda_j}, \ldots, e^{-i(n-1)\lambda_j})$, $\lambda_j = 2\pi j/n$, $j = 0, \ldots, n-1$. Also, let $w_{z_s} = W_s Z_{s,n}$, where W_s is defined as W, but for low frequency variables. As in Corbae et al. (2002), $BM(\Omega)$ denotes a vector Brownian motion with covariance matrix Ω and let the integrals $\int_0^1 B(r)dr$ be written as $\int_0^1 B$. $MN(0, G)$ represents a mixed normal distribution with matrix mixing variate G.

The paper is organized as: Section 2 discusses about FDC under regular datasets; Section 3 presents the asymptotic analysis of temporal aggregation; Section 4 introduces the proposed estimator; Section 5 refers to I(0) regressors efficiency; Section 6 presents the nonparametric analysis; Section 7 reports some Monte Carlo experiments; Section 8 presents an empirical application with quarterly GDP and monthly US indicators; Section 9 concludes the paper. Appendix 1.A presents the functional form of some matrices, Appendix 1.B presents lemmas and theorems proofs, and Appendix 1.C reports Monte Carlo simulation results.

2 Frequency dependent coefficient and MFD

It is commonplace in applied economic analyses to exclude periods of series related to crises, wars, or due to some unexpected outliers. The logical argument is that in these periods the series interact between them in a different pattern than previously observed. Engle (1974) argued that much effort was made to show the distinction between these periods, but just a small effort was directed to show the discrepancy of coefficients over frequencies. For example, under a crisis period we may have no variation between the relation of two macroeconomic series at low frequencies - long-run - but a variation at high frequencies - short-run - relation, or vice-versa. Under another premise, Corbae et al. (2002) used the concept of FDC to show that detrending in time domain could lead to inconsistent estimators.

Following Corbae et al. (2002) lines, we assume that the pair (e_t, x_t') satisfies one of the two subsequent sets of assumptions.

Assumption 1. *$\varsigma_t = (e_t, x_t')'$ is a jointly stationary time series with Wold representation $\varsigma_t = \sum_{j=0}^{\infty} C_j \xi_{t-j}$, where $\xi_t = iid(0, \Sigma)$ with finite fourth moments and the coefficients C_j satisfy $\sum_{j=0}^{\infty} j||C_j|| < \infty$. The spectral density matrix $f_{\varsigma\varsigma}(\lambda)$ of ς_t is defined by*

$$f_{\varsigma\varsigma}(\lambda) = \begin{bmatrix} f_{ee}(\lambda) & 0 \\ 0 & f_{xx}(\lambda) \end{bmatrix}$$

with $f_{ee}(\lambda), f_{xx}(\lambda) > 0 \ \forall \lambda$.

Assumption 2. *x_t is an I(1) process satisfying $\Delta x_t = v_t$, initialized at $t = 0$ by any $O_p(1)$ variable. The shocks $\varsigma_t = (e_t, v_t')'$ satisfy Assumption 1. The spectral density $f_{\varsigma\varsigma}(\lambda)$ of ς_t is defined by*

$$f_{\varsigma\varsigma}(\lambda) = \begin{bmatrix} f_{ee}(\lambda) & 0 \\ 0 & f_{vv}(\lambda) \end{bmatrix}$$

with $f_{ee}(\lambda), f_{vv}(\lambda) > 0 \ \forall \lambda$.

In a general framework, we can have for each frequency a possible distinct coefficient that drives the relationship between Y_n, $n \times 1$, and X_n, $n \times p$, which can be described as,

$$w_y(\omega) = w_x(\omega)\beta(\omega) + w_e(\omega), \quad \omega \in [-\pi, \pi]. \tag{1.1}$$

Now, let $\mathscr{B}_A = [-\omega_0, \omega_0]$ and $\mathscr{B}_A^c = [-\pi, -\omega_0) \cup (\omega_0, \pi]$, for some frequency $\omega_0 \in (0, \pi)$. Then, an interesting case of (1.1) is given by

$$w_y(\omega) = w_x(\omega)\mathbb{1}[\omega \in \mathscr{B}_A]\beta_A + w_x(\omega)\mathbb{1}[\omega \in \mathscr{B}_A^c]\beta_{A^c} + w_e(\omega), \quad \omega \in [-\pi, \pi], \tag{1.2}$$

where β_A and β_{A^c}, both $p \times 1$, represent the low and the high-frequency coefficient, or the long and the short run coefficients, respectively. Based on Corbae et al. (2002), this model represents the simplest model that contain frequency dependent coefficients, with just one coefficient for lower frequencies and other for higher frequencies.

For a finite number of observations, given a selector matrix A composed with ones on the main diagonal for those Fourier frequencies comprehended in $\mathscr{B}_A$ set, or A^c for $\mathscr{B}_A^c$ set, the Band Spectrum regression of the semiparametric model (1.2) is defined as

$$\hat{\beta}_A = [(\Psi X_n)'\Psi X_n]^{-1}(\Psi X_n)'\Psi Y_n, \quad \hat{\beta}_{A^c} = [(\Psi^c X_n)'\Psi^c X_n]^{-1}(\Psi^c X_n)'\Psi^c Y_n,$$

where prime means the transpose conjugate, $\Psi = AW$, $\Psi^c = A^c W$, and $AA^c = A^c A = 0$. Note that when $\beta_A = \beta_{A^c}$ in equation (1.2) we have a constant coefficient for all frequencies and there is no incentive for a treatment in the frequency domain.

In general, for equation (1.1) one can estimate nonparametrically the coefficients for a given frequency, say λ_k, using a narrow band selector matrix A_k via equation (1.3), where, as $n \to \infty$, $A_k W$ can be written as a shrinking operator in the neighborhood of frequency λ_k. For this, let the matrix $\Psi(\lambda_k) = A_k W$, $n \times n$, be the discrete Fourier transform around λ_k. Its $\{j,t\}$-element is described by $\varphi_{jt} = n^{-1/2}e^{i\lambda_j t}$, $t = 1,\ldots,n$, $\lambda_j = (2\pi j/n)$, $j = k - m^*,\ldots,k,\ldots,k+m^*$, $m = 2m^* + 1$, $1/m + m/n \to 0$, and 0 otherwise. Then, for $\lambda_k \to \omega$, $\omega \in [0, 2\pi]$, the estimate of $\beta(\omega)$ is defined as

$$\hat{\beta}(\omega) = [(\Psi(\lambda_k)X_n)'\Psi(\lambda_k)X_n]^{-1}(\Psi(\lambda_k)X_n)'\Psi(\lambda_k)Y_n. \tag{1.3}$$

3 Temporal Aggregation

Let y_t and x_t be two time series with a relationship in the frequency domain described by equation (1.2), where observations of x_t are obtained at a sampling frequency normalized to 1 observation per unit of time, and observations of y_t are only obtained at sampling frequency $1/s$, for some integer s. The subsampled series of y_t is designed as $y_{s,t}$ with $n_s = \lfloor n/s \rfloor$ observations. Henceforth, we assume that n is mod s, which imply that x_t has sn_s, observations.

Apart from series y_t being not fully observable, it is only for these specific periods $t = \{s, 2s, 3s, \ldots, n\}$. This implies not only a re-sample procedure but also a shift in time. Lemma 1 defines the relationship between y_t and $y_{s,t}$ in frequency domain.

Lemma 1. *Let $(w_y; w_{y_s})$ be the Fourier transform of $(Y_n; Y_{s,n})$, where $y_{s,t}$ is observed at $t = s, 2s, \ldots, n$; $n \mod s$. Then,*

$$\sqrt{s}w_{y_s}(\lambda_k) = \sum_{j=0}^{s-1} w_y(\lambda_k + 2\pi j/s)e^{-i(\lambda_k + 2\pi j/s)\cdot(s-1)}, \quad \forall \lambda_k \in \mathscr{B}_s,$$

where $2\omega_s = 1/s$ and $\mathscr{B}_s = [-\omega_s, \omega_s]$.

Lemma 1 relies on the fact that n mod s. For n not mod s, the folding process results in a not perfect alignment of the Fourier frequencies, requiring in a not straightforward spectral estimation without dropping observations. Furthermore, every sampling process implies a spectrum folding, which is not problematic for a sampling rate equal to or higher than the Nyquist rate. Nyquist-Shannon Sampling Theorem, see Shannon (1949), defines sufficient conditions for a signal to be properly sampled. Except in some circumstances, like data compression, it also imposes necessary conditions. A proper sampling occurs when the signal is sampled at a frequency, at least, two times higher than the highest frequency contained in the signal. If the Fourier coefficients of w_y are not equal to zero for frequencies higher than ω_s then w_{y_s} suffers from aliasing.

We can express Lemma 1 in matrix form as $F_s w_{y_s} = FDw_y$, where D, $n \times n$, is a diagonal delay matrix, F, $n \times n$, is a folding matrix, and F_s, $n \times n_s$, is an alignment matrix. See the Appendix A for the functional form of D, F, and F_s. Moreover, splitting the spectrum of w_y in two sets, say $\mathscr{B}_A$ and $\mathscr{B}_A^c$, we have $F_s w_{y_s} = FDAw_y + FDA^c w_y$.

Thereby, the least squares estimation, in frequency domain, of $Y_{s,n}$ on $X_{s,n}$, is

$$\tilde{\beta}_F = \left([(\Phi + \Phi^c)X_n]'[(\Phi + \Phi^c)X_n]\right)^{-1}[(\Phi + \Phi^c)X_n]'\Phi_s Y_{s,n}, \tag{1.4}$$

where $\Phi = FDAW$, $\Phi^c = FDA^cW$ and $\Phi_s = F_sW_s$. Before presenting the asymptotic behavior of $\tilde{\beta}_F$, see Theorem 1 and 2, let us present some definitions.

Definition 1. *Let $w_z(\lambda_k)$ be the Fourier transform of z_t, $t = 1, \ldots, n$, for frequency $\lambda_k = 2\pi k/n$, $k = 0, \ldots, n-1$, and let $w^d_{z_j}(\lambda_k)$ be defined as*

$$w^d_{z_j}(\lambda_k) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} z_t e^{-i(\lambda_k + 2\pi j/s)t} \cdot e^{-i(\lambda_k + 2\pi j/s)(s-1)} \mathbb{1}[\lambda_k \in \mathscr{B}_s],$$

where $j = 0, \ldots, s-1$. Assuming that the spectral density function of z_t exist, it is denoted as f_{zz} and the cross spectral density function of ${w_z}^d_j$ and ${w_z}^d_l$ is denoted by $f_{z^d_j z^d_l}$.

In terms of matrix F and D, $w^d_{z_j}$ can be written as $FD(w_z(\lambda)\mathbb{1}[\lambda \in \mathscr{B}^j_s])$, $\lambda \in [0, 2\pi]$, where $\mathscr{B}^j_s = [-\omega_s - 2\pi(j+1)/s, -\omega_s - 2\pi j/s) \cup (\omega_s + 2\pi j/s, \omega_s + 2\pi(j+1)/s]$.

Definition 2. *Let the pair (e_t, x'_t) satisfy Assumption 1. Then, as $n \to \infty$, we have*

(i) $\Sigma_A = \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x^d_j x^d_j}(\omega)\, \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A] d\omega$,

(ii) $\Sigma_{A^c} = \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x^d_j x^d_j}(\omega)\, \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}^c_A] d\omega$,

(iii) $\Sigma_F = \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega)\, d\omega$,

Taking results of Lemma 1 and Definition 1-3, we can state the inconsistency of temporal aggregation in the stationary series, Theorem 1, and the impossibility to recovery the short term coefficient in the nonstationary series, Theorem 2.

Theorem 1. *If (e_t, x_t') are zero mean, stationary and ergodic time series satisfying Assumption 1 and y_t is generated by equation (1.2), then the least squares estimation defined by (1.4) leads to a linear combination of β_A and β_{A^c}, given by*

$$\tilde{\beta}_F \xrightarrow{p} {\Sigma_F}^{-1}\Sigma_A\beta_A + {\Sigma_F}^{-1}\Sigma_{A^c}\beta_{A^c}.$$

Theorem 1 implies that $\tilde{\beta}_F$ is inconsistent whenever $\beta_A \neq \beta_{A^c}$ coefficients. For $\beta_{A^c} = \beta_A$, we have $\Sigma_F^{-1}(\Sigma_A + \Sigma_{A^c}) = 1$ which implies in the consistency of $\tilde{\beta}_F$.

For nonstationary series, we have,

Definition 3. *Let the pair (e_t, x_t') satisfy Assumption 2. Then*

(i) $\Theta_A = \int_0^1 B_x B_x'$,

(ii) $\Theta_{AU} = \int_0^1 B_x dB_u$.

Theorem 2. *If (e_t, x_t') satisfies Assumption 2 and y_t is generated by equation (1.2), then the least squares estimation defined by (1.4) leads to a consistent estimation of β_A coefficient, and*

$$n(\tilde{\beta}_F - \beta_A) \xrightarrow{d} MN\left(0, \Theta_A^{-1}\Theta_{AU}\right).$$

The temporal aggregation in the nonstationary case implies in the consistency of recovering the long-term coefficient and in the impossibility of estimating the short-term coefficient. The logic of this result is that, as x_t has long memory, the information contained in the spectrum becomes concentrated at low frequencies as n increases.

4 Aliased Band Spectrum Regression

In this section, we propose a novel method to handle distinct regression coefficients across frequencies in mixed-frequency datasets. Our procedure is the following: first, assuming that ω_0 is known, we split the spectrum of X as in equation (1.2); then, we multiply both sides of this equation by the folding matrix, F, together with the delay matrix, D, implying a folding process for each frequency band at the ω_s mark and in a rotation for each frequency by $e^{-i\lambda s}$. Equation (1.5) summarizes the procedure.

$$FDw_y \quad = \quad FDw_x^A\beta_A + FDw_x^{A^c}\beta_{A^c} + FDw_e. \tag{1.5}$$

Furthermore, by the Lemma 1, the LHS of equation (1.5) can be replaced by $F_s w_{y_s}$. As result we have

$$F_s w_{y_s} = FDw_x^A\beta_A + FDw_x^{A^c}\beta_{A^c} + w_u. \tag{1.6}$$

We call equation (1.6) the Aliased Band Spectrum, ABS, regression. Under (1.6), it is possible to regress the sub-sampled $y_{s,t}$ series against a regular sampled x_t series. Applying to both sides of equation (1.6) an inverse Fourier transform with delay, $W'D'$, and a sampling matrix S, which selects only the available observations of y_t series, results in

$$SW'D'F_s w_{y_s} \quad = \quad SW'D'FDw_x^A\beta_A + SW'D'FDw_x^{A^c}\beta_{A^c} + SW'D'w_u. \tag{1.7}$$

The LHS of equation (1.7) is the original $y_{s,t}$ series. The first and second term of RHS can be written as $x_{s,t}^A$ and $x_{s,t}^{A^c}$, respectively. Note that $x_{s,t}^{A^c} = x_{s,t} - x_{s,t}^A$, i.e., the re-sampled version of x_t at same sample frequency of $y_{s,t}$ minus the re-sampled and filtered x_t for frequencies belonging to the folded $\mathscr{B}_A$ set. Then, equation (1.7) can be rewritten in time domain as

$$y_{s,t} \quad = \quad x_{s,t}^A\beta_A + x_{s,t}^{A^c}\beta_{A^c} + \tilde{u}_{s,t}. \tag{1.8}$$

The sampling process, here described by matrix S, does not imply in any additional loss of information since the entire spectrum of x_t and y_t were allocated in the $\mathscr{B}_s$ set. Thus, equations (1.6) and (1.8) produce the same results, with errors in the frequency domain for the former and in the time domain for the latter.

Using $\Phi = FDAW$, $\Phi^c = FDA^cW$ and $\Phi_s = F_sW_s$, we define the ABS least squares estimates as

$$\begin{bmatrix} \tilde{\beta}_A \\ \tilde{\beta}_{A^c} \end{bmatrix} = \left([\Phi X_n \quad \Phi^c X_n]'[\Phi X_n \quad \Phi^c X_n]\right)^{-1}[\Phi X_n \quad \Phi^c X_n]'\Phi_s Y_{s,n}, \tag{1.9}$$

and its equivalent in time domain, with $\Xi = SW'D'$, as

$$\begin{bmatrix} \tilde{\beta}_A \\ \tilde{\beta}_{A^c} \end{bmatrix} = \left([\Xi\Phi X_n \quad \Xi\Phi^c X_n]' [\Xi\Phi X_n \quad \Xi\Phi^c X_n] \right)^{-1} [\Xi\Phi X_n \quad \Xi\Phi^c X_n]' Y_{s,n}.$$

Together with Lemma 4, Theorem 1' presents the consistency of the ABS estimator, equation (1.9), for stationary series.

Definition 4. *Let the pair (e_t, x_t') satisfy Assumption 1. Then, as $n \to \infty$, we have*

(i) $\Sigma_{AU} = 2\pi \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega) f_{uu}(\omega) \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A] d\omega,$

(ii) $\Sigma_{A^c U} = 2\pi \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega) f_{uu}(\omega) \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A^c] d\omega.$

Theorem 1'. *If (e_t, x_t') are zero mean, stationary and ergodic time series satisfying Assumption 1 and y_t is generated by equation (1.2). Regressing $y_{s,t}$ on x_t as in (1.9) leads to a consistent estimation of β_A and β_{A^c}. The joint limit distribution of $\tilde{\beta}_A$ and $\tilde{\beta}_{A^c}$ is given by*

$$\sqrt{n} \begin{pmatrix} \tilde{\beta}_A - \beta_A \\ \tilde{\beta}_{A^c} - \beta_{A^c} \end{pmatrix} \xrightarrow{d} N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_A^{-1} \Sigma_{AU} \Sigma_A^{-1} & 0 \\ 0 & \Sigma_{A^c}^{-1} \Sigma_{A^c U} \Sigma_{A^c}^{-1} \end{bmatrix} \right).$$

Lemma 5 and Theorem 2' present the consistency of ABS estimator, equation (1.9), for nonstationary series.

Definition 5. *Let the pair (e_t, x_t') satisfy Assumption 2. Then, as $n \to \infty$, we have*

(i) $\Theta_{A^c} = \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega) \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A^c] d\omega$
$\quad + B_x(1) B_x(1)' \sum_{j=0}^{s-1} \sum_{l=0}^{s-1} \int_{\mathscr{B}_s} g_{jl}(\omega) d\omega,$

(ii) $\Theta_{A^c U} = \sum_{j=0}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega) f_{uu}(\omega) \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A^c] d\omega$
$\quad + B_x(1) B_x(1)' \sum_{j=0}^{s-1} \sum_{l=0}^{s-1} \int_{\mathscr{B}_s} g_{jl}(\omega) f_{uu}(\omega) d\omega,$

(iii) $g_{jl}(\omega) = e^{-i(\omega + 2\pi(j-l)/s)} \left(\dfrac{\mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A^c] \mathbb{1}[\omega + 2\pi l/s \in \mathscr{B}_A^c]}{2\pi \left[1 - e^{-i(\omega + 2\pi j/s)}\right] \left[1 - e^{i(\omega + 2\pi l/s)}\right]} \right).$

Theorem 2'. *If (e_t, x_t') satisfies Assumption 2 and y_t is generated by equation (1.2), then ABS regression, (1.9), is consistent for β_A and β_{A^c}. The joint limit distribution of $\tilde{\beta}_A$ and $\tilde{\beta}_{A^c}$ is given by*

$$\Upsilon_n \left(\begin{array}{c} \tilde{\beta}_A - \beta_A \\ \tilde{\beta}_{A^c} - \beta_{A^c} \end{array} \right) \xrightarrow{d} MN\left(\left[\begin{array}{c} 0 \\ 0 \end{array} \right], \left[\begin{array}{cc} \Theta_A^{-1}\Theta_{AU} & 0 \\ 0 & \Theta_{A^c}^{-1}\Theta_{A^cU}\Theta_{A^c}^{-1} \end{array} \right] \right)$$

where $\Upsilon_n = diag(n^{-1}, n^{-1/2})$.

Empirically, one would desire to test if the coefficients $\tilde{\beta}_A$ and $\tilde{\beta}_{A^c}$ are statistically different from each other given a frequency ω_0 that split the sets $\mathscr{B}_A$ and $\mathscr{B}_A^c$. For this purpose, the residuals can be obtained by $w_{\hat{u}} = F_s w_{y_s} - FDw_x^A\tilde{\beta}_A - FDw_x^A\tilde{\beta}_{A^c}$. However, a main concern is the frequency zero behavior for both I(0) and I(1) spectra. For the former case, the spectrum at 0 represents the sample mean and the addition of a constant in the regression may circumvent eventual problems. However, in the latter, the spectrum is not bounded at such frequency. Taufemback (2018b) suggest to substitute $w_z(0)$ by $|w_z(1)|$. Another possibility is to trim frequency 0 from the spectrum. Both procedures have similar finite sample results for our problem. With frequency 0 spectrum controlled, the convergence described by Theorem 2' is now $\sqrt{n}$-consistent but with a non-diagonal asymptotic variance. Lemma 3, in Appendix, shows that the Fourier transform of an I(1) series is correlated with its frequencies.

The theorems of Section 3 and 4 rely on the spectrum division in only two sets, however, the underlining model may not be so restricted. The estimation for multiple FDC follows equation (1.9) lines, with selector matrices A_j, $j = 1, \ldots, b$, conditioned to $\mathscr{B}_j$, $j = 1, \ldots, b$ sets, and $\cup_{j=1}^b \mathscr{B}_j$ equals to the whole spectrum. However, the correct number of frequency-dependent coefficients and their respective frequency sets are generally unknown. Engle (1974) proposed an F-test to evaluate the presence of distinct coefficient for different band spectra. Furthermore, Engle (1980) proposed three different test statistics, Lagrange multiplier, Wald, and Log-likelihood ratio, that are robust in the presence of heteroskedasticity in the frequency domain, i.e., for non-white noise residuals.

These test formulations can be adapted to our purpose. Notice that, for a set of possible breaks in the spectrum relation we can follow the literature of structural breaks, see Andrews (1993); Bai and Perron (1998, 2003). The null hypothesis is such that the model has l FDC against $l+1$ FDC under the alternative. Thus, the test statistics for these three testing principles are,

$$\begin{aligned} \hat{\xi}_{LR} &= \textstyle\sum_{\lambda_k \in \mathscr{B}_s} \log(\bar{f}_{uu,l}(\lambda_k)/\bar{f}_{uu,l+1}(\lambda_k)) && (1.10) \\ \hat{\xi}_W &= (w_{\hat{u}_l}'\hat{\Omega}_{l+1}^{-1}w_{\hat{u}_l} - w_{\hat{u}_{l+1}}'\hat{\Omega}_{l+1}^{-1}w_{\hat{u}_{l+1}})/(w_{\hat{u}_{l+1}}'\hat{\Omega}_{l+1}^{-1}w_{\hat{u}_{l+1}}/n_s) && (1.11) \\ \hat{\xi}_{LM} &= (w_{\hat{u}_l}'\hat{\Omega}_l^{-1}w_{\hat{u}_l} - w_{\hat{u}_{l+1}}'\hat{\Omega}_l^{-1}w_{\hat{u}_{l+1}})/(w_{\hat{u}_l}'\hat{\Omega}_l^{-1}w_{\hat{u}_l}/n_s) && (1.12) \end{aligned}$$

with $\bar{f}_{uu,l}(\lambda_k) = m^{-1}\sum_{\lambda\in\mathscr{B}_k} w_{\hat{u}_l}(\lambda)w_{\hat{u}_l}(\lambda)'$, where $\mathscr{B}_k$ is a neighborhood around frequency λ_k, $\mathscr{B}_k = [\lambda_{k-m/2},\ldots,\lambda_{k+m/2}]$, $\hat{D}_l^{-1} = diag(\bar{f}_{uu,l}^{-1})$ for frequency belonging to $\mathscr{B}_s$ and 0 otherwise, and finally, $\hat{\Omega}_l^{-1} = W'\hat{D}_l^{-1}W$. Engle (1980) argues that all tests formulations are asymptotically distributed as χ_p^2 under the null. His argumentations can be extended to our case, based on ABS regressions, without further discussions. In Section 7, we present an extensive Monte Carlo experiment covering the finite properties of these tests.

5 Efficiency of stationary ABS regression

It is well known that efficiency declines whether the error term of equation (1.2) in the time domain is non-spherical, or non-flat in the frequency domain. However, these error patterns can be corrected by GLS techniques in the frequency domain, see Hannan (1963a). Robinson (1991a) delimited necessary conditions for convergence of the weighted - or smoothed - periodogram to the spectral density matrix f_{uu}, where $w_{\hat{u}}$ are the residuals of a stationary series regression.Assume that the even smoothing function K satisfies,

$$\int_R |K(\lambda)|d\lambda < \infty, \quad \int_R K(\lambda)d\lambda = 1, \quad k_t = \int_R K(\lambda)e^{it\lambda}d\lambda, \tag{1.13}$$

where R is the real line, and K and the bandwidth M respect assumptions A2(j) and A3(ν) of Robinson (1991a), transcribed below.

Assumption A2(j): K is real and even and satisfies (1.13); k_t satisfies $|k_t| \leq \bar{k}_t$, where

$$\int_0^\infty (1+t^j)\bar{k}_t dt < \infty,$$

and $\bar{k}_t$ is monotonically decreasing on $[0,\infty)$ and chosen to be even.

Assumption A3(ν): $M^{-1} + Mn^{-\nu} \xrightarrow{p} 0$, as $n\to\infty$.

The convergence of $\hat{f}_{uu}$ only requires $j = 0$ in Assumption A2(j), but equal to 2 to satisfy GLS necessary conditions. The parameter ν must lie on $(0, 1/2]$ range. Hence, let the periodogram of $w_{\hat{u}}$ be defined by,

$$I_{uu}(\lambda) = w_{\hat{u}}(\lambda)w_{\hat{u}}(\lambda)', \quad \text{for } \lambda \in \mathscr{B}_s,$$

therefore, let $K_M(\lambda)$ represents the kernel K conditioned to the bandwidth M, we have the estimated spectral density given by

$$\hat{f}_{uu}(\lambda) \;=\; \begin{cases} \dfrac{2\pi M}{n} \displaystyle\sum_{j \in \mathscr{B}_s} K_M(\lambda - \bar{\lambda}_j) I_{uu}(\bar{\lambda}_j) & , \text{for} \quad \lambda \in \mathscr{B}_s \\ 0 & , \text{for} \quad \lambda \notin \mathscr{B}_s \end{cases}$$

and $\bar{\lambda} \in \mathcal{B}_s$, where $\mathcal{B}_s$ is a set composed by the periodic repetition of $\mathscr{B}_s$ set. The Aliased Band Spectrum regression using generalized least squares estimated, ABS-GLS, is defined by equation (1.14) and Theorem 3 defines its asymptotic properties.

$$\begin{bmatrix} \tilde{\beta}_A^{GLS} \\ \tilde{\beta}_{A^c}^{GLS} \end{bmatrix} = \Big([\Phi X_n \quad \Phi^c X_n]' \hat{\Omega}^+ [\Phi X_n \quad \Phi^c X_n] \Big)^{-1} [\Phi X_n \quad \Phi^c X_n]' \hat{\Omega}^+ \Phi_s Y_{s,n} \tag{1.14}$$

where $\hat{\Omega}^+ = diag[\hat{f}_{uu}^{-1}(\lambda_k) \mathbb{1}[\lambda_k \in \mathscr{B}_s]]$. Notice that, the inverse of $\hat{f}_{uu}(\lambda)$ only is required for $\lambda \in \mathscr{B}_s$, as the values of Ω^+ are assumed to be 0 for frequencies on the complementary set.

Definition 6. *Let the pair (e_t, x_t') satisfy Assumption 1. Then, as $n \to \infty$,*

$$\begin{array}{lcl} (i) \quad \Sigma_{A/u} & = & (2\pi)^{-1} \sum_{j=1}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega) f_{uu}^{-1}(\omega) \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A] d\omega, \\ (ii) \quad \Sigma_{A^c/u} & = & (2\pi)^{-1} \sum_{j=1}^{s-1} \int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega) f_{uu}^{-1}(\omega) \mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A^c] d\omega. \end{array}$$

Theorem 3. *If (e_t, x_t') are zero mean, stationary and ergodic time series that satisfy Assumption 1, Assumption A2(2), Assumption A3(ν), $\nu \in (0, 1/2]$, and y_t is generated by equation (1.2), then the efficient ABS regression is consistent for β_A^{GLS} and $\beta_{A^c}^{GLS}$. The joint limit distribution of $\tilde{\beta}_A^{GLS}$ and $\tilde{\beta}_{A^c}^{GLS}$ is given by*

$$\sqrt{n} \begin{pmatrix} \tilde{\beta}_A^{GLS} - \beta_A \\ \tilde{\beta}_{A^c}^{GLS} - \beta_{A^c} \end{pmatrix} \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, V_{GLS} \right),$$

where,

$$V_{GLS} = \begin{bmatrix} \Sigma_{A/u} & 0 \\ 0 & \Sigma_{A^c/u} \end{bmatrix}^{-1} \qquad \text{and} \quad V \geq V_{GLS}.$$

Whenever u_t presents temporal dependency, f_{uu} varies across frequencies implying that $V > V_{GLS}$. In case that $f_{uu}(\omega)$ equals to a constant for all $\omega \in [0, 2\pi]$, $V = V_{GLS}$.

6 Nonparametric Analysis

The nonparametric model, see Corbae et al. (2002), described by

$$y_t = \sum_{j=-\infty}^{\infty} \beta_j' x_{t-j} + e_t = \beta(L)' x_t + e_t \tag{1.15}$$

has no restriction in respect the behavior of coefficients across frequencies, and also does not censor the presence of leads. In frequency domain, it is assumed that $b(\omega) = \beta(e^{i\omega}) = \sum_{j=-\infty}^{\infty} \beta_j e^{ij\omega}$ converges for all $\omega \in [-\pi, \pi]$. Furthermore, applying the Beveridge-Nelson decomposition and the discrete Fourier transform, we have for every frequency $\lambda_k = 2\pi k/n$ under Assumption 1,

$$w_y(\lambda_k) = b(\lambda_k)' w_x(\lambda_k) + w_e(\lambda_k) + O_p(n^{-1/2}) \tag{1.16}$$

and under Assumption 2,

$$w_y(\lambda_k) = \beta(1)' w_x(\lambda_k) - \bar{\beta}(e^{i\lambda_k})' w_v(\lambda_k) + w_e(\lambda_k) + O_p(n^{-1/2}) \tag{1.17}$$

with $b(-\omega) = \beta(1) + \bar{\beta}(e^{i\omega})(e^{i\omega} - 1)$ for $\omega \neq 0$.

Due the frequency superposition caused by the subsampling process, we need to proceed as in Section 4. Thus, to estimate nonparametrically $b(-\lambda_k)$, $\lambda_k \in \mathscr{B}_s$, we need to define $\Phi_j^k = FDI_j^k W$, $j = 0, \ldots, s-1$, where I_j^k is a selector matrix in a neighborhood of frequency $\lambda_k + 2\pi j/s$, $\lambda \in [\lambda_{k-m}, \lambda_{k+m}]$, with $1/m + m/n \to 0$ as $n \to \infty$. Also, $\Phi_s^k = F_s I_s^k W_s$, where I_s^k is a selector matrix with dimension $n_s \times n_s$.

For I(0) and I(1), let $\check{\beta}$ and $\tilde{\beta}$ indicate the narrow band estimates using temporal aggregation and ABS regression, respectively. The results present in this section are, in general, similar to those present in Sections 3 and 4. For I(0) series, temporal aggregation procedure implies in a general inconsistent estimator and ABS procedure is able to recovery all coefficients. Theorems 4 and 4', allied with Definition 7, present the results for the stationary case.

Definition 7. *Let the pair (e_t, x_t') satisfy Assumption 1, and as $n \to \infty$, $\lambda_k \to \omega$, then*

(i) $\Lambda_F(\omega) = 2\pi \sum_{j=1}^{s-1} f_{x_j^d x_j^d}(\omega)$,

(ii) $\Lambda_j(\omega) = 2\pi f_{x_j^d x_j^d}(\omega)$,

(iii) $\Lambda_{jU}(\omega) = (2\pi)^2 f_{x_j^d x_j^d}(\omega) f_{uu}(\omega)$,

Theorem 4. *If (e_t, x_t') are zero mean, stationary and ergodic time series satisfying Assumption 1 and y_t is generated by equation (1.15), then for $\forall j = 0, \dots, s-1$, the narrow band estimation on temporal aggregated (1.16) leads to a linear combination of coefficients given by*

$$\check{b}(\lambda_k) \quad \xrightarrow{p} \quad \sum_{j=0}^{s-1} \Lambda_F^{-1}(\omega)\Lambda_j(\omega)b(\omega + 2\pi j/s).$$

Similar to the semiparametric case, see Section 3, whether $b(\omega) = b(\omega + 2\pi j/s)$ for all $j \in \{0, \dots, s-1\}$, the estimator $\check{b}(\lambda_k)$ is consistent.

Theorem 4'. *If (ε_t, x_t') are zero mean, stationary and ergodic time series satisfying Assumption 1 and y_t is generated by equation (1.15). Then the ABS regression on (1.16) leads to consistent estimates of $b(\omega + 2\pi j/s), j = 0, \dots, s-1$. The joint limit distribution is given by*

$$\sqrt{n} \begin{pmatrix} \tilde{b}(\omega) - b(\omega) \\ \vdots \\ \tilde{b}(\omega + 2\pi(s-1)/s) - b(\omega + 2\pi(s-1)/s) \end{pmatrix} \xrightarrow{d} N(0, V_{4'}(\omega))$$

where

$$V_{4'}(\omega) = \begin{bmatrix} \Lambda_0^{-1}(\omega)\Lambda_{0U}(\omega)\Lambda_0^{-1}(\omega) & & 0 \\ & \ddots & \\ 0 & & \Lambda_{(s-1)}^{-1}(\omega)\Lambda_{(s-1)U}(\omega)\Lambda_{(s-1)}^{-1}(\omega) \end{bmatrix}.$$

For nonstationary series, we need to analyze two main cases, one with ω equal to 0 and other with ω distinct of 0. For $\omega \neq 0$ we can re-write the spectrum of x_t in terms of its first difference $\hat{v}_t$. Thus, the equation (1.17) in terms of v_t becomes,

$$w_y(\lambda_k) = \left[\frac{w_v(\lambda_k)}{1 - e^{i\lambda_k}} + o_p(1)\right]\beta(1) + \left[w_v(\lambda_k) + o_p(1)\right]\beta(e^{i\lambda_k}) + \tilde{E}(\lambda_k) \tag{1.18}$$

with

$$b_\omega = \beta(1)(1 - e^{i\lambda_k})^{-1} + \beta(e^{i\lambda_k}) \text{ and } b_\omega(1 - e^{i\lambda_k}) \rightarrow b(-\omega) \text{ as } n \rightarrow \infty,$$

according to Corbae et al. (2002). Notice that for $\omega = 0$ and for the ABS regression on (1.17), we have as regressors the vector $[w_x(\lambda), w_v(\lambda + 2\pi/s), \dots, w_v(\lambda + 2\pi(s-1)/s)]$, $\lambda \in \mathscr{B}_0$. Thus, the asymptotic behavior of estimators for the nonstationary case is given by Definition 8, Theorems 5, and 5'.

Definition 8. *Let the pair (e_t, x_t') satisfies Assumption 2. Then, as $n \to \infty$, we have*

(i) $f_{x_j^d x_j^d}(\omega) = |1 - e^{i(\omega + 2\pi j/s)}|^{-2} f_{v_j^d v_j^d}(\omega)$,

(ii) $\Gamma_F(\omega) = \sum_{j=0}^{s-1} f_{x_j^d x_j^d}(\omega)$.

Theorem 5. *If (e_t, x_t') satisfies Assumption 2 and y_t is generated by equation (1.15), then*

(a) if $\omega = 0$, the narrow band regression on temporal aggregated (1.17) leads to a consistent estimation of the long run coefficient.

$$n(\check{b}(0) - \beta(1)) \xrightarrow{d} \left(\int_0^1 B_x B_x' \right)^{-1} \int_0^1 B_x dB_u$$

(b) if $\omega \neq 0$, the narrow band regression on temporal aggregated (1.18) leads to a linear combination of coefficients given by

$$\check{b}(\omega) \quad \xrightarrow{p} \quad \sum_{j=0}^{s-1} \Gamma_F^{-1}(\omega) f_{x_j^d x_j^d}(\omega) b(\omega + 2\pi j/s)$$

Theorem 5'. *If (e_t, x_t') satisfies Assumption 2 and y_t is generated by equation (1.15), then*

(a) if $\omega = 0$, the augmented ABS regression on (1.17) leads to the consistency of $\tilde{b}(2\pi j/s)$, $j = 0, \ldots, s-1$. The joint limit distribution is given by

$$\Upsilon_m \begin{pmatrix} \tilde{b}(0) - \beta(1) \\ \tilde{b}(2\pi/s) - b_{2\pi/s} \\ \vdots \\ \tilde{b}(2\pi(s-1)/s) - b_{2\pi(s-1)/s} \end{pmatrix} \xrightarrow{d} MN(0, V_{5'}(0))$$

where,

$$V_{5'}(0) = \begin{bmatrix} \left(\int_0^1 B_x B_x' \right)^{-1} \int_0^1 B_x dB_u & & & 0 \\ & f_{uu}(0) f^{-1}_{x_1^d x_1^d}(0) & & \\ & & \ddots & \\ 0 & & & f_{uu}(0) f^{-1}_{x_{s-1}^d x_{s-1}^d}(0) \end{bmatrix}$$

and Υ_m is a $s \times s$ matrix equal to $diag(n, n^{1/2}, \ldots, n^{1/2})$.

(b) if $\omega \neq 0$, the augmented ABS regression on (1.18) leads to a consistent estimation of $\tilde{b}(\omega + 2\pi j/s), j = 0, \ldots, s-1$. The joint limit distribution is given by

$$\sqrt{n}\begin{pmatrix} \tilde{b}(\omega) - b_\omega \\ \vdots \\ \tilde{b}(\omega + 2\pi(s-1)/s) - b_{\omega+2\pi(s-1)/s} \end{pmatrix} \xrightarrow{d} N(0, V_{5'}(\omega))$$

where

$$V_{5'}(\omega) = \begin{bmatrix} f_{uu}(\omega) f^{-1}_{x_1^d x_1^d}(\omega) & & 0 \\ & \ddots & \\ 0 & & f_{uu}(\omega) f^{-1}_{x_{s-1}^d x_{s-1}^d}(\omega) \end{bmatrix}.$$

7 Finite sample properties

Macroeconomic series are, generally, reported at levels. Assuming that they are integrated at order one, first differentiation without missing data give us the desired stationary series. In case of quarterly GDP, we have in levels the sum of the last three monthly GDP observations, say $\bar{y}_t^q = \bar{y}_t^m + \bar{y}_{t-1}^m + \bar{y}_{t-2}^m$; where letters with bars represent variables in levels. For example, first quarter GDP means the sum of January, February, and March. The sum of December, January, and February, if available, would represent $\bar{y}_{t-1}^q$. The difference between $\bar{y}_t^q$ and $\bar{y}_{t-1}^q$ give us the monthly innovation of $\bar{y}_t^q$ series, or its growth rate if we take log differences. However, GDP comes every quarter with non-overlapping months. Differentiate from a quarter apart brings up some conjunction of shocks in a pattern given by equation (1.19). Assuming, $\bar{y}_t^m = \bar{y}_0^m + \sum_{j=0}^{t-1} y_{t-j}^m$, and Δ_q be the differentiation of two consecutive quarters, then,

$$\begin{aligned} \Delta_q \bar{y}_{s,t}^q &= \bar{y}_t + \bar{y}_{t-1} + \bar{y}_{t-2} - (\bar{y}_{t-3} + \bar{y}_{t-4} + \bar{y}_{t-5}), \\ &= \sum_{j=0}^{t} y_{t-j}^m + \sum_{j=0}^{t-1} y_{t-j}^m + \sum_{j=0}^{t-2} y_{t-j}^m - \sum_{j=0}^{t-3} y_{t-j}^m - \sum_{j=0}^{t-4} y_{t-j}^m - \sum_{j=0}^{t-5} y_{t-j}^m \qquad (1.19) \\ &= y_t^m + 2y_{t-1}^m + 3y_{t-2}^m + 2y_{t-3}^m + y_{t-4}^m. \end{aligned}$$

Consequently, to regress $\Delta_q \bar{y}_{s,t}$ on monthly indicators, x_t, we need to reproduce the same pattern of aggregation on monthly series, say $\Delta_q \bar{x}_t = x_t + 2x_{t-1} + 3x_{t-2} +$

$2x_{t-3} + x_{t-4}$. Alternatively, one can generate a new series representing the sum of last three months at levels and then differentiate. For the nonstationary series, we can sum the last three months, $\bar{x}_t^q = \bar{x}_t + \bar{x}_{t-1} + \bar{x}_{t-2}$, to reproduce the quarterly GDP aggregation. Mariano and Murasawa (2003, 2010) suggest the use of the geometric mean instead of the accounting identity, as we use. Since they use log difference to transform the series from levels to stationarity, the geometric mean represents a good approximation and linear steady-state models can be employed. Despite the geometric mean for $\Delta_q \log y_{s,t}$ be a good approximation for the unobserved $\Delta \log y_t$. We stick with the accounting identity method to prevent us from introducing other errors, which could be an issue on tests performance.

Therefore, to stress the reliability of the tests presented in Section 4, we ran a Monte Carlo experiment to measure the rejection frequency of the null hypothesis, l vs $l+1$ FDC. Using a sequential procedure we tested up to 8 FDC, with $\alpha = 0.01$. The Monte Carlo experiment follows,

$$\begin{aligned} y_t &= \textstyle\sum_{l=1}^{q} (W' A_l W) x_t \beta_l + \sigma e_t, \\ z_t &= \theta z_{t-1} + u_t, \; u_t \sim W.N.(0,1), \\ e_t &= \rho e_{t-1} + v_t, \; v_t \sim W.N.(0,1), \end{aligned} \tag{1.20}$$

where x_t was generated in two fashions for, both, I(0) and I(1). First as systematic skip sampling stock variables and second as a simulation of the GDP sampling scheme. For I(0), we have $x_t = z_t$ and $x_t = z_t + 2z_{t-1} + 3z_{t-2} + 2z_{t-3} + z_{t-4}$. For I(1), we have $x_t = z_t$ and $x_t = z_t + z_{t-1} + z_{t-2}$. The series y_t is subsampled at one observation for every three observations of x_t. For I(0) and I(1), $\rho = 0.3$ and $\sigma = 0.5$. For I(0), $\theta = 0.8$ and $\theta = 1$ for I(1). The maximum number of FDC was set as 5, with structural breaks in the spectrum randomly assigned between $[0.10\pi, 0.80\pi]$ with minimal distance between two consecutive breaks of 0.10π. The values of β_l are independently randomly assigned on the set $[-2, 2]$, but with $\min |\beta_l - \beta_{l+1}| > 0.5$.

Like in Andrews (1993); Bai and Perron (1998, 2003), the breakpoints candidates were selected using a minimization of SSR. We also avoid the first h_i and the last h_l frequencies. As well, we defined as h_b the minimal number of frequencies in a segment. Where $h_k = 0.10$, for $k = i, b, l$. The sample sizes chosen were $\{80, 120, 160\}$ for the low-frequency variable, corresponding to $\{20, 30, 40\}$ years of data in a monthly-quarterly sample scheme. Notice that, due to the spectrum symmetry, for a large number of observations the values of h_i and h_l can be reduced to close to 0.

As expected, all three tests statistics performed similarly, see Tables C.1-1 to C.1-12 of Appendix C, and comparable with time-domain methods for structural break, see Bai and Perron (1998). The results behave similarly for I(0) and (1), as well as for the distinct sample procedures. However, whenever we work with a time domain series with structural breaks, we expect to have roughly the same signal-to-noise ratio, SNR, over time. However, in the frequency domain, whether the model presents some temporal dependency, the spectrum of dependent variable will present distinct SNR across the frequencies. For example, series following 'the typical spectrum shape of an economic variable' have the majority of its variability explained by its low-frequency spectrum, see Granger (1966a). For these series, the detection of a break in high frequencies could be compromised. Because, due to unfavorable SNR, regressions in this region are consistent but inefficient when compared with regressions in low frequencies. In our experiment, as the number of FDC increases also increases the probability of having more breaks at high frequencies. Resulting in a tendency, for all cases, to select the true model less frequently as q increases.

Finally, we explore the use of AIC, $\ln(SSR)+2(q-1)p/n_s$, and BIC, $\ln(SSR)+\ln(n_s)(q-1)p/n_s$, as an alternative selector of the number of different slopes. The results, see Tables C.1-13 to C.1-16, are in Appendix C. AIC presented unsatisfactory results, selecting in general more breaks than the correct model presents. Thus, AIC results are not reported. BIC had comparable results with the three test presented above. In fact, in experiments not reported, as the number of variables in x_t, p, increases BIC selects the true model more frequently. The other tests present similar results as p increases.

8 Frequency relation among US indicators

Stock and Watson (1989, 1991) introduced an alternative coincident index for US economy extracting a common factor of four monthly coincident series. Mariano and Murasawa (2003) extended their work with MFD by introducing quarterly GDP in a similar framework - state space representation allied with Kalman filter. Notwithstanding, in Mariano and Murasawa (2010) the authors argued that a coincident index is, in fact, a substitute of a monthly GDP. Thus, estimating an indicator of the monthly GDP results to be more attractive since it provides an economic interpretation.

Here, we exploit the same set of variables from Mariano and Murasawa (2010), see Table 1-1, to investigate the existence of possible breaks in the frequency relation of US indicators. The set of variables, except by quarterly GDP, it is the same set of Stock and Watson (1989). In fact, Stock and Watson (1989) uses employees-hours instead of the total number of employees. However, the US Conference Board, reports every month a coincident index based on the four monthly series reported in Table 1-1.

Following the guidelines of Section 7, we set (h_i, h_b, h_l) as (0.10,0.10,0.20), i.e., we avoid the initial 10% and the last 20% of the frequencies, and with a minimal space between two consecutive breaks of 10% of the number of frequencies between

Table 1-1: US indicators from January of 1959 to June of 2015.

Description
Quarterly
Real GDP[1] (billions of chained 2009 dollars, SA, AR).
Monthly
Employees on non-agricultural payrolls[2] (thousands, SA);
Personal income less transfer payments[3] (bi. of chained 2009 dollars, SA, AR);
Index of industrial production[4] (2012=100, SA);
Manufacturing and trade sales[3] (millions of chained 2009 dollars, SA).

Notes: SA means 'seasonally adjusted' and AR means 'annual rate'. Source: (1) BEA - Bureau of Economic Analysis, U.S. Department of Commerce, (2) Bureau of Labor Statistics, U.S. Department of Labor, (3) The Conference Board, (4) Federal Reserve, United States.

$[0, \pi]$. Then, using variables in levels and first differences, we tested up to 5 FDC with no lags for neither the exogenous nor the endogenous variables.

Our findings indicate that the reconstruction of a monthly GDP will be biased under temporal aggregation, since all methods, present in Table 1-2, rejected the presence of a unique coefficient for all frequencies. In fact, BIC selects one break, 2 FDC, for variables in levels and first differences. As mention in Section 7, AIC tends to report more FDC than the correct model. Here, AIC reported at least 5 FDC for stationary series and 3 FDC for nonstationary. Following the literature of structural break test, see Bai and Perron (1998, 2003), the sequential test is repeated until the test statistic indicates no further rejection of the null hypothesis. Thus, for variables in level, all three tests present in Section 4 also indicate 2 FDC. Finally, for variables in first difference, Wald and Lagrange Multiplier test report the presence of 3 FDC and Log-likelihood test indicate 4 FDC. Notice that the list of breakpoint candidates is similar for I(0) and I(1) variables.

9 Conclusion

The MFD literature has plenty of empirical papers demonstrating the importance of the intermediate data. Aside from this work, no paper had addressed the asymptotic effects of mixed-sampling data under frequency-dependent coefficients. Understanding how sub-sampling in time domain affects the series spectrum is the cornerstone of our analysis. Here, we showed that splitting the spectrum of the independent variable in frequency bands, and mimicking the same folding

Table 1-2: Test results for conditional presence, l vs $l+1$, of frequency dependent coefficients on the linear relationship between monthly US indicators and quarterly GDP.

	H_0	$\hat{\xi}_W$	$\hat{\xi}_{LR}$	$\hat{\xi}_{LM}$	AIC	BIC	$\bar{\lambda}$
I(0)	1 vs 2	**0.032**	**0.000**	**0.011**	5.588	5.588	0.28π
	2 vs 3	**0.000**	**0.008**	**0.000**	5.544	**5.632**	0.46π
	3 vs 4	0.274	**0.015**	0.394	5.483	5.659	0.16π
	4 vs 5	0.798	0.381	0.811	5.451	5.715	0.80π
	5 vs 6	0.283	0.397	0.304	**5.448**	5.800	0.66π
	H_0	$\hat{\xi}_W$	$\hat{\xi}_{LR}$	$\hat{\xi}_{LM}$	AIC	BIC	$\bar{\lambda}$
I(1)	1 vs 2	**0.000**	**0.000**	**0.000**	18.014	18.014	0.10π
	2 vs 3	0.906	1.000	0.885	17.926	**18.013**	0.58π
	3 vs 4	0.000	0.000	0.000	**17.911**	18.086	0.69π
	4 vs 5	0.967	0.067	0.962	17.805	18.068	0.80π
	5 vs 6	0.973	0.844	0.986	17.812	18.162	0.48π

Notes: Bold values for $\hat{\xi}_W$, $\hat{\xi}_{LR}$, and $\hat{\xi}_{LM}$ tests represent p-values below 5%. Bold values for AIC and BIC indicate the minimum value for each model criteria. $\bar{\lambda}$ indicate break points candidates. For covariance matrix estimation, m was chosen as $n_s^{0.25}$, see Section 4.

process that happened with the dependent variable, any linear relationship can be consistently recovered. Also, as described in Section 3, temporal aggregation leads to problems in the coefficients estimation.

The ABS method imposes milder conditions when compared with others mixed-frequency methods present in the literature, as well it relies on the condition that all series are generated at the same frequency. For example, interpolation methods rely on smooth transitions between observables and unobservables low-frequency values, see Harvey and Pierse (1984) and Bernanke et al. (1997). ABS regression does not require it. In fact, the presence of higher frequencies than the Nyquist frequency limit is a key feature of our method. MIDAS methods make use of non-usual DGPs with different frequencies driving the generate process, see Clements and Galvão (2008), Marcellino and Schumacher (2010), Foroni et al. (2015b), Guérin and Marcellino (2013). Also, MIDAS relies on non-linear regression for many of its parametrizations. Moreover, ABS regression provides access to the band spectrum relationships between series sampled at different frequencies, bringing up a new set of analysis of macroeconomic series. Efficient computation and simple implementation are also characteristics that must be highlighted.

Furthermore here, we study the presence of multiple frequency dependent coefficients in the linear relationship between US monthly indicators and GDP. Our findings show a rejection of a single coefficient for all frequencies. Which indicate that a reconstruction of monthly GDP using temporal aggregation may be affected by a coefficient bias. Recently, McCracken and Ng (2016) presented an extensive compendium of monthly, and some quarterly, series from St. Louis Federal Reserve Economic Data (FRED) database. Thus, as future work, one could explore the frequency domain relationship between the series comprehended in this large dataset.

Appendix 1.A

The functional form of D, F, and F_s is given by: D, $n \times n$, is a diagonal matrix where $d_{kk} = e^{-i\lambda_k(s-1)}$, for $\lambda_k = 2\pi k/n$ with $k = 0, \ldots, n-1$ and

$$F = \left[\begin{array}{c:c:c:c} IO_\tau & IO_\tau & \ldots & IO_\tau \\ \hdashline \multicolumn{4}{c}{O_{n-n_s \times n}} \\ \hdashline IO_{\tau^*} & \ldots & IO_{\tau^*} & IO_{\tau^*} \end{array}\right]_{n \times n}, \quad F_s = \sqrt{s} \left[\begin{array}{c:c} I_{\tau \times \tau} & O_{\tau \times n_s - \tau} \\ \hdashline \multicolumn{2}{c}{O_{n-n_s \times n_s}} \\ \hdashline O_{\tau^* \times n_s - \tau^*} & I_{\tau^* \times \tau^*} \end{array}\right]_{n \times n_s},$$

where $IO_\tau = [I_{\tau\times\tau} \quad O_{\tau\times\tau^*}]$, $IO_{\tau^*} = [O_{\tau^*\times\tau} \quad I_{\tau^*\times\tau^*}]$, $\tau = \lfloor n_s/2 \rfloor + 1$ and $\tau^* = n_s - \tau$.

Appendix 1.B

Proof. **Lemma 1:** Following the proof of Theorem 4 and 9 of Cooley et al. (1969), let $z_t, t = 0, \ldots, n-1$ be a sequence of natural numbers evenly spaced and

$$z_t = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} w_z(\lambda_k) e^{i2\pi kt/n}, \tag{1.21}$$

$$w_z(\lambda_k) = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} z_t e^{-i2\pi kt/n}. \tag{1.22}$$

Now suppose a sampling process over z_t, taking one sample at every s observations, $n \mod s$. Then for $t' = 0, \ldots, n/s - 1$ and taking first observation at $t = s$, then equation (1.21) becomes,

$$\begin{aligned} z_{st'-\Delta_t} &= \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} w_z(\lambda_k) e^{i2\pi k(st'-\Delta_t)/n} \\ &= \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} w_z(\lambda_k) e^{-i2\pi k\Delta_t/n} e^{i2\pi kt'/n_s}, \end{aligned} \tag{1.23}$$

where $n_s = n/s$. Using (1.22) in (1.23) for $k^* \in \mathscr{B}_s$, i.e., $k^* \in [0, n_s - 1]$,

$$\begin{aligned} w_z(\lambda_{k^*}) &= \frac{1}{\sqrt{n_s}} \sum_{t'=0}^{n_s-1} \left(\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} w_z(\lambda_k) e^{-i2\pi k\Delta_t/n} e^{i2\pi kt'/n_s} \right) e^{-i2\pi k^* t'/n_s} \\ &= \frac{\sqrt{s}}{n} \sum_{t'=0}^{n_s-1} \sum_{k'=0}^{n_s-1} \left[\sum_{j=0}^{s-1} w_z[k' + jn_s] e^{-i2\pi(k'+jn_s)\Delta_t/n} e^{i2\pi(k'+jn_s)t'/n_s} \right] e^{-i2\pi k^* t'/n_s} \\ &= \frac{\sqrt{s}}{n} \sum_{t'=0}^{n_s-1} \left[\sum_{j=0}^{s-1} w_z(\lambda_k^* + jn_s) e^{-i2\pi(k^*+jn_s)\Delta_t/n} \right] \\ &= \frac{\sqrt{s}}{n} \frac{n}{s} \sum_{j=0}^{s-1} w_z(\lambda_k^* + jn_s) e^{-i2\pi(k^*+jn_s)\Delta_t/n} \\ &= \frac{1}{\sqrt{s}} \sum_{j=0}^{s-1} w_z(\lambda_k^* + jn_s) e^{-i2\pi(k^*+jn_s)\Delta_t/n}. \end{aligned}$$

□

Lemma 2. *If (e_t, x_t') are zero mean, stationary and ergodic time series satisfying Assumption 1 and Definition 1, then as $n \to \infty$, $(\lambda_k + 2\pi j/s) \to \omega + 2\pi j/s$, $j = 0, \ldots, s-1$, then*

$$
\begin{aligned}
(i) \qquad & n_s^{-1}(\Phi X_n)'\Phi X_n &=& \; n_s^{-1}\textstyle\sum_{\lambda_k \in \mathscr{B}_s}\sum_{j=1}^{s-1}\sum_{l=1}^{s-1} w_{x_j}'^{d} w_{x_l}^{d}\mathbb{1}[\lambda_k \in \mathscr{B}_A] \\
& &\xrightarrow{p}& \; \Sigma_A, \\
(ii) \qquad & n_s^{-1}(\Phi^c X_n)'\Phi^c X_n &=& \; n_s^{-1}\textstyle\sum_{\lambda_k \in \mathscr{B}_s}\sum_{j=1}^{s-1}\sum_{l=1}^{s-1} w_{x_j}'^{d} w_{x_l}^{d}\mathbb{1}[\lambda_k \in \mathscr{B}_A^c] \\
& &\xrightarrow{p}& \; \Sigma_{A^c}, \\
(iii) \quad & n_s^{-1/2}[\Phi X_n \quad \Phi^c X_n]' w_{\hat{u}} &\xrightarrow{d}& \; N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{AU} & 0 \\ 0 & \Sigma_{A^cU} \end{bmatrix}\right)
\end{aligned}
$$

Proof. **Lemma 2:** The proofs follow closely Corbae et al. (1997) and Corbae et al. (2002). Let $\#\{\lambda_k \in \mathscr{B}_A\}$ denote the number of fundamental frequencies in the band $\mathscr{B}_s$. Subdividing $[-\pi, \pi]$ in small and equal sub-bands $\mathscr{B}_b$ with π/B length, each one centered on frequencies $\{\omega_b = \pi b/B : b = -B+1, \ldots, B-1\}$. Let $m = \#(\lambda_k \in \mathscr{B}_b)$, then $n = 2mB$ and $n_{\mathscr{B}_A} = 2mB_s$, approximately. Assuming that $m, B \to \infty$ and $B/\sqrt{n} \to 0$ as $n \to \infty$, we fulfill Assumption 3 of Corbae et al. (2002).

Thus under Assumption 1, we have

$$
\begin{aligned}
n_s^{-1}\sum_{\lambda_k \in \mathscr{B}_s} w_{x_j}^d(\lambda_k) w_{x_l}^d(\lambda_k)' &= n_s^{-1}\textstyle\sum_{b=-B_s+1}^{B_s-1}\sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_{x_l}^d(\lambda_k)' \\
&= \tfrac{1}{2B}\textstyle\sum_{b=-B_s+1}^{B_s-1}\tfrac{1}{m}\sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_{x_l}^d(\lambda_k)' + o_p(1) \\
&= \tfrac{1}{2B}\textstyle\sum_{b=-B_s+1}^{B_s-1} 2\pi \hat{f}_{x_j^d x_l^d}(\omega_b) \\
&\xrightarrow{p} \textstyle\int_{\mathscr{B}_s} f_{x_j^d x_l^d}(\omega)d\omega > 0 \text{ if } j = l,
\end{aligned} \tag{1.24}
$$

where $\omega_b = \pi b/B$ and since $\hat{f}_{x_j^d x_j^d}(\omega_b) = (2\pi m)^{-1}\sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_{x_j}^d(\lambda_k)'$ is a consistent estimator of $f_{x_j^d x_j^d}(\omega)$. For $j \neq l$, the cross product is equal to zero due frequency orthogonality. Using these results, (i)-(ii) follows.

For the convergence in distribution, by Corbae et al. (2002) we have,

$$n_s^{-1/2} w_{x_j}^d w_{\hat{u}} = n^{-1/2} \sum_{\lambda_k \in \mathscr{B}_s} w_{x_j}^d(\lambda_k) w_u(\lambda_k)' = n_s^{-1/2} \sum_{b=-B_a+1}^{B_a-1} \sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_u(\lambda_k)'. \tag{1.25}$$

Notice that $w_u(\lambda_k)$ satisfies a CLT for discrete Fourier transforms, see Hannan (1970) and Hannan (1973) for a central limit theorem for discrete Fourier transform, i.e., $m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_j} w_u(\lambda_k) \xrightarrow{d} N(0, 2\pi f_{uu}(\omega))$, and by Assumption 1, the spectrum of (1.25) is asymptotically independent, thus for $\lambda_k \in \mathscr{B}_b$, and

$$m^{-1} \sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_{x_j}^d(\lambda_k)' \xrightarrow{p} 2\pi f_{x_j^d x_j^d}(\omega)$$

and

$$\begin{aligned} n_s^{-1/2} \sum_{b=-B_a+1}^{B_a-1} \sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_u(\lambda_k)^* &= \frac{1}{\sqrt{2B}} \sum_{b=-B_a+1}^{Ba-1} \frac{1}{\sqrt{m}} \sum_{\lambda_k \in \mathscr{B}_b} w_{x_j}^d(\lambda_k) w_u(\lambda_k)^* \\ &\xrightarrow{d} N\left(0, 2\pi \int_{\mathscr{B}_A} f_{x_j^d x_j^d}(\omega) f_{uu}(\omega) d\omega\right) \end{aligned} \tag{1.26}$$

with similar result for $\mathscr{B}_A^c$ set.

Finally, we need to find the limit variance of $[\Phi X_n \quad \Phi^c X_n]' w_u$. Gathering all previous steps,

$$\begin{aligned} \mathbb{V}\left(n_s^{-1/2}[\Phi X_n \quad \Phi^c X_n]' w_u\right) &= n_s^{-1} \mathbb{E}\left([\Phi X_n \quad \Phi^c X_n]' w_u w_u' [\Phi X_n \quad \Phi^c X_n]\right) \\ &= n_s^{-1} \mathbb{E} \begin{bmatrix} (\Phi X_n)' w_u w_u' \Phi X_n & (\Phi X_n)' w_u w_u' \Phi^c X_n \\ (\Phi X_n)' w_u w_u' \Phi^c X_n & (\Phi^c X_n)' w_u w_u' \Phi^c X_n \end{bmatrix}, \end{aligned}$$

by independence of x_t and e_t, we have,

$$\begin{aligned} \mathbb{V}\left(n_s^{-1/2}[\Phi X_n \quad \Phi^c X_n]' w_u\right) &= n_s^{-1} \begin{bmatrix} \mathbb{E}[(\Phi X_n)' \Phi X_n] \mathbb{E}[w_u w_u'] & \mathbb{E}[(\Phi X_n)' \Phi^c X_n] \mathbb{E}[w_u w_u'] \\ \mathbb{E}[(\Phi X_n)' \Phi^c X_n] \mathbb{E}[w_u w_u'] & \mathbb{E}[(\Phi^c X_n)' \Phi^c X_n] \mathbb{E}[w_u w_u'] \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_{n,AU} & \Sigma_{n,AA^cU} \\ \Sigma_{n,AA^cU} & \Sigma_{n,\tilde{A}U} \end{bmatrix} \end{aligned}$$

and using the results of (1.26) we have,

$$
\begin{aligned}
\Sigma_{n,AU} &= n_s^{-1}\sum_{\lambda_k\in\mathscr{B}_s}\sum_{j=0}^{s-1}\sum_{l=0}^{s-1}\mathbb{E}[w^d_{x_j}(\lambda_k)w^d_{x_l}(\lambda_k)']\mathbb{E}[w_u(\lambda_k)w_u(\lambda_k)']\mathbb{1}[\lambda_k\in\mathscr{B}_A]\\
&= \Sigma_{AU}+o(1),\\
\Sigma_{n,A^cU} &= n_s^{-1}\sum_{\lambda_k\in\mathscr{B}_s}\sum_{j=0}^{s-1}\sum_{l=0}^{s-1}\mathbb{E}[w^d_{x_j}(\lambda_k)w^d_{x_l}(\lambda_k)']\mathbb{E}[w_u(\lambda_k)w_u(\lambda_k)']\mathbb{1}[\lambda_k\in\mathscr{B}_A^c]\\
&= \Sigma_{A^cU}+o(1),\\
\Sigma_{n,AA^cU} &= n_s^{-1}\sum_{\lambda_k\in\mathscr{B}_A}\sum_{j=0}^{s-1}\sum_{l=0}^{s-1}\mathbb{E}[w^d_{x_j}(\lambda_k)w^d_{x_j}(\lambda_k)']\mathbb{E}[w_u(\lambda_k)w_u(\lambda_k)']\\
&= o(1),\quad \text{by the orthogonality of discrete Fourier transform.}
\end{aligned}
$$

Since $\mathbb{E}[X^d_j(\lambda_k),X^d_l(\lambda_k)]=0$, for $j\neq l$, one can understand as two independent series and use the results of Corbae et al. (2002).

□

Proof. **Theorem 1:** We have,

$$
\begin{aligned}
\tilde\beta_F &= M_1^{-1}w_{z_1},\ \text{where}\\
M_1 &= (\Phi X_n)'(\Phi X_n)+(\Phi^c X_n)'(\Phi^c X_n)+2(\Phi X_n)'(\Phi^c X_n),\ \text{and}\\
w_{z_1} &= [(\Phi X_n)'(\Phi X_n)'+(\Phi X_n)'(\Phi^c X_n)]\beta_A\\
&+ [(\Phi^c X_n)'(\Phi^c X_n)'+(\Phi X_n)'(\Phi^c X_n)]\beta_{A^c}+w_u,
\end{aligned}
$$

notice that by Assumption 1,

$$
n_s^{-1}[(\Phi+\Phi^c)x]'w_{\hat u} = n_s^{-1}\sum_{i=0}^{n-1}\sum_{l=0}^{s-1}w^d_{x_l}(\lambda_i)\sum_{j=0}^{s-1}w_{e_j}(\lambda_i)\xrightarrow{p}0,
$$

Then for $\beta=\beta_A=\beta_{A^c}$,

$$
\tilde\beta_F = M_1^{-1}\left([(\Phi X_n)'(\Phi X_n)+(\Phi^c X_n)'(\Phi^c X_n)]\beta\right)+o_p(1)\xrightarrow{p}\beta
$$

notice that $M_1^{-1}>0$ by (1.24).

The limit variance of $[(\Phi+\tilde{\Phi})x_t]'w_u$ is given by

$$\begin{aligned}
\mathbb{V}\left(n_s^{-1/2}[(\Phi+\Phi^c)X_n]'w_u\right) &= n_s^{-1}\mathbb{E}\left([(\Phi+\Phi^c)X_n]'w_u w_u'[(\Phi+\Phi^c)X_n]\right) \\
&= n_s^{-1}\sum_{j=0}^{s-1}\sum_{\lambda_k\in\mathscr{B}_s}\mathbb{E}\left(w_{x_j}^d(\lambda_k)w_{x_j}^d(\lambda_k)'\right)\mathbb{E}\Big(w_u(\lambda_k)w_u(\lambda_k)'\Big) \\
&= 2\pi\sum_{j=0}^{s-1}\int_{\mathscr{B}_s} f_{x_j^d x_j^d}(\omega)\, f_{uu}(\omega)\, d\omega + o(1).
\end{aligned}$$

where the joint convergence is given by (1.26), and

$$\begin{aligned}
\Sigma_{n,F} &= n_s^{-1}\sum_{j=0}^{s-1}\sum_{\lambda_k\in\mathscr{B}_s}\mathbb{E}\Big(w_{x_j}^d(\lambda_k)w_{x_j}^d(\lambda_k)'\Big) \\
&= \Sigma_F + o(1)
\end{aligned} \tag{1.27}$$

We have,

$$\begin{aligned}
\tilde{\beta}_F &= [(\Phi X_n)'(\Phi X_n)+(\Phi^c X_n)'(\Phi^c X_n)]^{-1}[(\Phi X_n)'(\Phi X_n)']\beta_A \\
&\quad +[(\Phi X_n)'(\Phi X_n)+(\Phi^c X_n)'(\Phi^c X_n)]^{-1}[(\Phi^c X_n)'(\Phi^c X_n)']\beta_{A^c} + o_p(1)
\end{aligned}$$

so that,

$$\tilde{\beta}_F \xrightarrow{p} {\Sigma_F}^{-1}\Sigma_A\beta_A + {\Sigma_F}^{-1}\Sigma_{A^c}\beta_{A^c}$$

by Lemma 2 and equation (1.27). □

Lemma 3. *If the pair (u_t, z_t') satisfies Assumption 2, then for $2\pi(k/n+j/s)\neq 0$, the Fourier transform of z_t is given by*

$$w_{z_j}^d(\lambda_k) = \left\{\frac{1}{1-e^{-i2\pi(k/n+j/s)}}V_j^d(\lambda_k) - \frac{e^{-i2\pi(k/n+j/s)}}{1-e^{-i2\pi(k/n+j/s)}}\frac{(z_n-z_0)}{\sqrt{n}}\right\}\mathbb{1}[\lambda_k\in\mathscr{B}_s]$$

for $j=0,\dots,s-1$ and $k=0,\dots,n-1$.

Proof. Lemma 3: The proof follows the lines of Lemma B of Corbae et al. (2002) and Definition 1. □

Lemma 4. *If the pair (e_t, x_t') satisfies Assumption 2, then as $n \to \infty$,*

$$
\begin{aligned}
(i) \quad & n_s^{-2}(\Phi X_n)'\Phi X_n &=& \; n_s^{-1}\textstyle\sum_{j=0}^{s-1}\sum_{l=0}^{s-1} w_{x_j}'^{d}\mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A]\times \\
& & & \; w_{x_l}^{d}\mathbb{1}[\omega + 2\pi l/s \in \mathscr{B}_A] \quad \xrightarrow{d} \quad \Theta_A \\
(ii) \quad & n_s^{-1}(\Phi^c X_n)'\Phi^c X_n &=& \; n_s^{-1}\textstyle\sum_{j=0}^{s-1}\sum_{l=0}^{s-1} w_{x_j}'^{d}\mathbb{1}[\omega + 2\pi j/s \in \mathscr{B}_A^c]\times \\
& & & \; w_{x_l}^{d}\mathbb{1}[\omega + 2\pi l/s \in \mathscr{B}_A^c] \quad \xrightarrow{d} \quad \Theta_{A^c} \\
(iii) \quad & n_s^{-3/2}(\Phi X_n)'\Phi^c X_n &\xrightarrow{p}& \; 0 \\
(iv) \quad & \Upsilon[\Phi X_n \quad \Phi^c X_n]' w_{\hat{u}} &\xrightarrow{d}& \; MN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Theta_{AU} & 0 \\ 0 & \Theta_{A^cU} \end{bmatrix}\right)
\end{aligned}
$$

Proof. Lemma 4:

- (i) Lemma C(a) of Corbae et al. (2002);
- (ii) Lemma C(e) of Corbae et al. (2002) and Lemma 3;
- (iii) With intention to simplify the proof and the notation, we omit here, w.l.o.g., the delay matrix D and also defined $\lambda_{j,k} = \lambda_k + 2\pi j/s$. We want to show that,

$$n^{-3/2}\sum_{\lambda_k \in \mathscr{B}_s}\left(\sum_j w_x(\lambda_{j,k})\mathbb{1}[\lambda_{j,k} \in \mathscr{B}_A]\right)'\left(\sum_j w_x(\lambda_{j,k})\mathbb{1}[\lambda_{j,k} \in \mathscr{B}_A^c]\right) \xrightarrow{p} 0$$

for $\lambda_k \neq 0$ the result is straightforward, by the $n^{-3/2}$ normalization factor and by the fact that there is no frequency with a singularity point. For $\lambda_k = 0$, let us assume that $\exists c > 0$ s.t. $2\pi j/s \in \mathscr{B}_A$, $j = 0, \dots, c-1$, and $2\pi j/s \in \mathscr{B}_A^c$, $j = c, \dots, s-1$. Then,

$$n_s^{-3/2}(w_x(0) + \cdots + w_x(2\pi(c-1)/s))'(w_x(2\pi c/s)) + \cdots + w_x(2\pi(s-1)/s))$$

we can rewrite the above expression using results for $\lambda_k \neq 0$, as

$$n_s^{-3/2}X(0)'(w_x(2\pi c/s)) + \cdots + w_x(2\pi(s-1)/s)) + o_p(1)$$

using Lemma 3 we have,

$$
\begin{aligned}
& n_s^{-3/2}\sum_{j=c}^{s-1} w_x(0)'\left[\frac{1}{1-e^{-i2\pi j/s}}w_{v_j}(2\pi j/s) - \frac{e^{-i2\pi j/s}}{1-e^{-i2\pi j/s}}\frac{(x_n - x_0)}{\sqrt{n}}\right] \\
&= n^{-1/2}s^{3/2}\sum_{j=c}^{s-1}\left[\frac{\sum_t x_t}{n^{3/2}}C_{j,1}C_{j,2} - C_{j,3}\frac{\sum_t x_t}{n^{3/2}}\frac{\sum_t v_t}{\sqrt{n}} + \frac{\sum_t x_t}{n^{3/2}}x_0\right]
\end{aligned}
$$

where $C_{j,1} = n^{-1/2}\sum_t v_t e^{-i2\pi jt/s}$, $C_{j,2} = 1/(1-e^{-i2\pi j/s})$, $C_{j,3} = e^{-i2\pi j/s}/(1-e^{-i2\pi j/s})$. By Phillips (1987) and Hamilton (1994), we know that $n^{-3/2}\sum_t x_t$ and $n^{-1/2}\sum_t v_t$ are stochastically bounded. Given the remaining factor of $n^{-1/2}$, we conclude our proof.

(*iv*) Following lines of Corbae et al. (2002) Theorem 3 proof and Lemma 4*(iii)*, the result follows.

□

Proof. **Theorem 2:** Using the algebra of proof of Theorem 1, and results of Lemma 4, it is easy to show that whenever $\beta = \beta_A = \beta_{A^c}$ we have $\check{\beta}_F \xrightarrow{p} \beta$. For $\beta_A \neq \beta_{A^c}$, note that

$$\begin{aligned} n_s^{-1}[(\Phi+\Phi^c)x]'w_{\hat{u}} &= n_s^{-1}\sum_{\lambda_k\in\mathscr{B}_s}\sum_{j=0}^{s-1} X_j^d U \\ &= n_s^{-1}\sum_{\lambda_k\in[-\pi,\pi]} X_0^d w_{\hat{u}} + o_p(1) \xrightarrow{d} \int_0^1 B_x dB_u. \end{aligned}$$

and by Lemma 4

$$n_s^{-2}(\Phi+\Phi^c)X_n = n_s^{-2}(\Phi)x + o_p(1) \quad \xrightarrow{d} \quad \int_0^1 B_x B_x',$$

thus,

$$\begin{aligned} \check{\beta}_F \;=\; & [(\Phi X_n)'(\Phi X_n) + (\Phi^c X_n)'(\Phi^c X_n) + 2(\Phi X_n)'(\Phi^c X_n)]^{-1}[(\Phi X_n)'(\Phi X_n)' + (\Phi X_n)'(\Phi^c X_n)]\beta \\ & [(\Phi X_n)'(\Phi X_n) + (\Phi^c X_n)'(\Phi^c X_n) + 2(\Phi X_n)'(\Phi^c X_n)]^{-1}[(\Phi^c X_n)'(\Phi^c X_n)' + (\Phi X_n)'(\Phi^c X_n) \end{aligned}$$

so that,

$$\check{\beta}_F \;=\; {\Theta_F}^{-1}\Theta_A\beta_A + o_p(1) \xrightarrow{p} \beta_A.$$

□

Proof. **Theorem 1':** From equation (1.9) and (1.6), we have,

$$\begin{bmatrix} \tilde{\beta}_A \\ \tilde{\beta}_{A^c} \end{bmatrix} = M_{1'}^{-1} w_{z_1'},$$

where,

$$\begin{aligned} M_{1'} &= \begin{pmatrix} (\Phi X_n)'\Phi X_n & (\Phi X_n)'\Phi^c X_n \\ (\Phi^c X_n)'\Phi X_n & (\Phi^c X_n)'\Phi^c X_n \end{pmatrix}, \\ w_{z_1'} &= \begin{pmatrix} (\Phi X_n)'\Phi X_n\beta_A + (\Phi X_n)'\Phi^c X_n\beta_{A^c} + (\Phi X_n)'E \\ (\Phi^c X_n)'\Phi X_n\beta_A + (\Phi^c X_n)'\Phi^c X_n\beta_{A^c} + (\Phi^c X_n)'E \end{pmatrix}. \end{aligned}$$

By Assumption 1, $n^{-1}(\Phi X_n)'U \xrightarrow{p} 0$ and $n^{-1}(\Phi^c x)'U \xrightarrow{p} 0$. Thus, we have, $\tilde{\beta}_A \xrightarrow{p} \beta_A$ and $\tilde{\beta}_{A^c} \xrightarrow{p} \beta_{A^c}$. For the asymptotic distribution we have,

$$\sqrt{n}\left(\begin{bmatrix} \tilde{\beta}_A - \beta_A \\ \tilde{\beta}_{A^c} - \beta_{A^c} \end{bmatrix}\right) = \{n^{-1}M_1\}^{-1}\cdot\{n^{-1/2}[\Phi X_n \quad \tilde{\Phi}x]'U\},$$

along the same lines of Lemma 2, the result follows. □

Proof. **Theorem 2':** Using the same algebra from proof of Theorem 1' and Lemma 4 result follows. □

Proof. **Theorem 3:** We can rely our proofs showing that our assumptions satisfy Robinson (1991a) assumptions, consequently allowing us to use his results. Thus, as described on section 5, K_M respects assumptions A2(j) and A3(ν) of Robinson (1991a). Furthermore, Assumption 1 fulfill requirements of Assumption A(ν), $\nu \in (0, 1/2]$, as well as Assumptions C1-C6. Consequently, Theorem 2.1 of Robinson (1991a) proves the convergence of $\hat{f}_{uu}(\lambda)$ to $f_{uu}(\lambda)$, for $\lambda \in [0, 2\pi]$. Theorems 4.1 and 4.2 proves that equation (1.14) results in a spherical normal random variable.

Taking $G = [\Phi X_n \quad \Phi^c x]$ and $\beta_{GLS} = [\beta_A^{GLS} \quad \beta_{A^c}^{GLS}]'$, and with a little bit of algebra equation (1.14) results in

$$\tilde{\beta}_{GLS} - \beta = (G'\hat{\Omega}^+ G)^{-1} G'\hat{\Omega}^+ U, \tag{1.28}$$

where RHS of equation (1.28) is $o(1)$ by Assumption 1 and Robinson's theorems. Notice that by equation (1.8), the two exogenous series can be treat as independent series, then by Theorem 4.2 of Robinson (1991a) and Theorem 1' proof, the asymptotic variance of RHS of equation 1.28 converges to V_{GLS}. Finally, to show that $V \geq V_{GLS}$, we need to show that $(G'G)^{-1}G'\Omega G(G'G)^{-1} - (G'\Omega^{-1}G)^{-1}$, or $(G'\Omega^{-1}G) - (G'G)(G'\Omega G)^{-1}(G'G)$ is positive semidefinite. Since, Ω is positive semidefinite by definition, taking $P = G'\Omega^{-1} - (G'G)(G'\Omega G)^{-1}G'$ then

$$P\Omega P' = (G'\Omega^{-1}G) - (G'G)(G'\Omega G)^{-1}(G'G) \geq 0$$

□

Lemma 5. *If (e_t, x_t') are zero mean, stationary and ergodic time series satisfying Assumption 1, then as $n \to \infty$,*

(i) $m^{-1} \sum_{\lambda_k \in \mathscr{B}_\omega} \sum_{j=0}^{s-1} \sum_{l=1}^{s-1} w_{x_j}'^{d}(\omega) w_{x_l}^{d}(\omega) \xrightarrow{p} \Lambda_F = 2\pi \sum_{j=1}^{s-1} f_{x_j^d x_j^d}(\omega)$,

(ii) $m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_\omega} \sum_{j=0}^{s-1} w_{x_j}'^{d}(\lambda_k) w_u(\lambda_k) \xrightarrow{d} \Lambda_{FU} = (2\pi)^2 \sum_{j=1}^{s-1} f_{x_j^d x_j^d}(\omega) f_{uu}(\omega)$,

(iii) $m^{-1} \sum_{\lambda_k \in \mathscr{B}_\omega} w_{x_j}'^{d}(\omega) X_j^d(\omega) \xrightarrow{p} \Lambda_j = 2\pi f_{x_j^d x_j^d}(\omega)$,

(iv) $m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_\omega} w_{x_j}'^{d}(\lambda_k) w_u(\lambda_k) \xrightarrow{d} \Lambda_{FU} = (2\pi)^2 f_{x_j^d x_j^d}(\omega) f_{uu}(\omega)$,

Proof. **Lemma 5:** Proof follows close to proof of Lemma 2 and of Corbae et al. (2002) Theorem 2, and it is omitted. □

Proof. **Theorem 4:** Using results of Lemma 5 and Theorem 1 proof, result follows. □

Proof. **Theorem 4':** Using results of Lemma 5 and Theorem 1' proof, result follows. □

Lemma 6. *If the pair (e_t, x_t') satisfies Assumption 2, $\Delta x_t = v_t$, then as $n \to \infty$,*

$$(i)\ m^{-1} \sum_{\lambda_k \in \mathscr{B}_\omega} \sum_{j=0}^{s-1} \sum_{l=1}^{s-1} w_{x_j}'^d(\lambda_k) w_{x_l}^d(\lambda_k) \xrightarrow{p} 2\pi \sum_{j=0}^{s-1} \frac{f_{v_j^d v_j^d}(\omega)}{|1 - e^{-i(\omega + 2\pi j/s)}|^2},$$

$$(ii)\ m^{-2} \sum_{\lambda_k \in \mathscr{B}_0} \sum_{j=0}^{s-1} \sum_{l=1}^{s-1} w_{x_j}'^d(\lambda_k) w_{x_l}^d(\lambda_k) \xrightarrow{d} \int_0^1 B_x B_x'$$

$$(iii)\ m^{-1} \sum_{\lambda_k \in \mathscr{B}_0} \sum_{j=0}^{s-1} \sum_{l=1}^{s-1} w_{x_j}'^d(\lambda_k) w_u(\lambda_k) \xrightarrow{d} \int_0^1 B_x dB_u$$

$$(iv)\ m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_0} \sum_{j=0}^{s-1} \sum_{l=0}^{s-1} V_j'^d(\lambda_k) V_l^d(\lambda_k) \xrightarrow{p} 2\pi \sum_{j=0}^{s-1} f_{v_j^d v_j^d}(0),$$

$$(v)\ m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_0} \sum_{j=0}^{s-1} V_j'^d(\lambda_k) w_u(\lambda_k) \xrightarrow{d} (2\pi)^2 \sum_{j=0}^{s-1} f_{v_j^d v_j^d}(0) f_{uu}(0),$$

$$(vi)\ m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_\omega} \sum_{j=0}^{s-1} \sum_{l=0}^{s-1} V_j'^d(\lambda_k) V_l^d(\lambda_k) \xrightarrow{p} 2\pi \sum_{j=0}^{s-1} f_{v_j^d v_j^d}(\omega),$$

$$(vii)\ m^{-1/2} \sum_{\lambda_k \in \mathscr{B}_\omega} \sum_{j=0}^{s-1} V_j'^d(\lambda_k) w_u(\lambda_k) \xrightarrow{d} (2\pi)^2 \sum_{j=0}^{s-1} f_{v_j^d v_j^d}(\omega) f_{uu}(\omega),$$

Proof. **Lemma 6:**

(i) See proof of Corbae et al. (2002) Theorem 2'b and 3'.

(ii) See proof of Corbae et al. (2002) Lemma C(j).

(iii) See proof of Corbae et al. (2002) Theorem 3'.

(iv) Note that for a stationary time series z_t,

$$m^{-1} \sum_{\lambda_k \in \mathscr{B}_\omega} w_z'(\lambda_k) w_z(\lambda_k) \xrightarrow{p} 2\pi f_{z_j^d z_j^d}(\omega)$$

thus, together with Lemma 5*(ii)* proof result follows.

(v) See proof of Corbae et al. (2002) Theorem 1'.

(vi) Same lines of Lemma 6(*iii*).

(vii) By independence of x_t and u_t, and Lemma 6(*iii*).

□

Proof. **Theorem 5:** Using results of Lemma 6 and Theorem 2 proof, result follows. □

Proof. **Theorem 5':** Using results of Lemma 6 and Theorem 2' proof, result follows. □

Appendix 1.C

Table C.1-1: Simulated Log-Likelihood, $\hat{\xi}_{LR}$, sequential test rejection statistics up to 8 FDC, see equation (1.10), for I(0) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.859	0.121	0.017	0.004	0.000	0.000	0.000	0.000
120	0.859	0.121	0.018	0.001	0.000	0.000	0.000	0.000
160	0.854	0.122	0.022	0.003	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.043	0.812	0.128	0.016	0.001	0.000	0.000	0.000
120	0.027	0.828	0.128	0.016	0.001	0.000	0.000	0.000
160	0.017	0.829	0.134	0.018	0.002	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.143	0.165	0.579	0.101	0.012	0.000	0.000	0.000
120	0.111	0.121	0.640	0.118	0.010	0.001	0.000	0.000
160	0.092	0.101	0.659	0.132	0.016	0.001	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.081	0.310	0.213	0.339	0.052	0.005	0.000	0.000
120	0.041	0.286	0.230	0.376	0.060	0.007	0.000	0.000
160	0.017	0.252	0.184	0.458	0.084	0.005	0.000	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.113	0.169	0.383	0.196	0.127	0.012	0.001	0.000
120	0.056	0.101	0.337	0.199	0.271	0.036	0.001	0.000
160	0.046	0.072	0.307	0.210	0.323	0.040	0.001	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-2: Simulated Log-Likelihood, $\hat{\xi}_{LR}$, sequential test rejection statistics up to 8 FDC, see equation (1.10), for I(0) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.813	0.158	0.026	0.002	0.000	0.000	0.000	0.000
120	0.803	0.172	0.023	0.003	0.000	0.000	0.000	0.000
160	0.796	0.180	0.022	0.002	0.001	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.044	0.790	0.147	0.017	0.001	0.000	0.000	0.000
120	0.028	0.793	0.158	0.021	0.001	0.000	0.000	0.000
160	0.015	0.787	0.176	0.021	0.001	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.114	0.148	0.614	0.115	0.010	0.000	0.000	0.000
120	0.114	0.123	0.645	0.108	0.009	0.001	0.000	0.000
160	0.110	0.084	0.664	0.128	0.012	0.000	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.051	0.240	0.233	0.411	0.062	0.004	0.000	0.000
120	0.032	0.224	0.207	0.455	0.075	0.005	0.001	0.000
160	0.030	0.183	0.229	0.472	0.078	0.007	0.000	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.048	0.088	0.352	0.215	0.262	0.033	0.002	0.000
120	0.044	0.065	0.384	0.230	0.246	0.030	0.000	0.000
160	0.031	0.052	0.405	0.190	0.274	0.046	0.002	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-3: Simulated Wald, $\hat{\xi}_W$, sequential test rejection statistics up to 8 FDC, see equation (1.11), for I(0) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.841	0.135	0.022	0.002	0.000	0.000	0.000	0.000
120	0.840	0.136	0.022	0.002	0.000	0.000	0.000	0.000
160	0.846	0.128	0.023	0.002	0.001	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.046	0.792	0.144	0.017	0.001	0.000	0.000	0.000
120	0.031	0.803	0.142	0.021	0.003	0.000	0.000	0.000
160	0.018	0.815	0.144	0.021	0.001	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.124	0.160	0.582	0.118	0.015	0.001	0.000	0.000
120	0.092	0.132	0.630	0.130	0.015	0.001	0.000	0.000
160	0.056	0.101	0.676	0.151	0.016	0.001	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.068	0.284	0.219	0.349	0.072	0.008	0.000	0.000
120	0.034	0.251	0.247	0.385	0.075	0.008	0.001	0.000
160	0.012	0.220	0.189	0.477	0.092	0.008	0.001	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.098	0.165	0.376	0.198	0.146	0.016	0.001	0.000
120	0.039	0.101	0.317	0.212	0.285	0.044	0.003	0.000
160	0.033	0.078	0.293	0.208	0.339	0.046	0.002	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-4: Simulated Wald, $\hat{\xi}_W$ sequential test rejection statistics up to 8 FDC, see equation (1.11), for I(0) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.804	0.163	0.029	0.004	0.000	0.000	0.000	0.000
120	0.799	0.168	0.028	0.004	0.001	0.000	0.000	0.000
160	0.800	0.166	0.029	0.005	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.048	0.771	0.154	0.024	0.003	0.000	0.000	0.000
120	0.029	0.781	0.162	0.026	0.003	0.000	0.000	0.000
160	0.014	0.772	0.183	0.028	0.002	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.080	0.158	0.610	0.132	0.019	0.002	0.000	0.000
120	0.096	0.127	0.637	0.123	0.016	0.001	0.000	0.000
160	0.086	0.087	0.673	0.136	0.016	0.002	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.026	0.225	0.242	0.428	0.072	0.006	0.000	0.000
120	0.017	0.208	0.216	0.462	0.087	0.010	0.000	0.000
160	0.014	0.174	0.234	0.480	0.086	0.012	0.000	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.020	0.069	0.356	0.233	0.275	0.042	0.004	0.000
120	0.019	0.056	0.381	0.243	0.259	0.040	0.002	0.000
160	0.015	0.042	0.399	0.191	0.298	0.053	0.002	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-5: Simulated Lagrange Multiplier, $\hat{\xi}_{LM}$, sequential test rejection statistics up to 8 FDC, see equation (1.12), for I(0) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.871	0.108	0.018	0.002	0.000	0.000	0.000	0.000
120	0.860	0.122	0.016	0.001	0.000	0.000	0.000	0.000
160	0.860	0.119	0.018	0.002	0.001	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.053	0.823	0.113	0.011	0.000	0.000	0.000	0.000
120	0.034	0.824	0.123	0.016	0.002	0.000	0.000	0.000
160	0.020	0.830	0.130	0.018	0.002	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.136	0.178	0.580	0.095	0.011	0.000	0.000	0.000
120	0.099	0.140	0.637	0.114	0.011	0.000	0.000	0.000
160	0.059	0.105	0.683	0.140	0.013	0.001	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.079	0.304	0.225	0.328	0.058	0.006	0.000	0.000
120	0.037	0.269	0.255	0.373	0.059	0.007	0.000	0.000
160	0.015	0.234	0.196	0.469	0.081	0.006	0.000	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.112	0.184	0.378	0.192	0.121	0.011	0.001	0.000
120	0.045	0.109	0.332	0.209	0.267	0.036	0.001	0.000
160	0.036	0.085	0.301	0.210	0.326	0.040	0.002	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-6: Simulated Lagrange Multiplier, $\hat{\xi}_{LM}$, sequential test rejection statistics up to 8 FDC, see equation (1.12), for I(0) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.835	0.143	0.020	0.002	0.000	0.000	0.000	0.000
120	0.827	0.148	0.023	0.002	0.000	0.000	0.000	0.000
160	0.821	0.151	0.024	0.004	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.056	0.795	0.131	0.016	0.002	0.000	0.000	0.000
120	0.033	0.801	0.148	0.017	0.002	0.000	0.000	0.000
160	0.016	0.797	0.165	0.021	0.001	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.097	0.168	0.619	0.104	0.011	0.001	0.000	0.000
120	0.107	0.132	0.644	0.103	0.013	0.001	0.000	0.000
160	0.093	0.091	0.680	0.121	0.014	0.001	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.030	0.253	0.246	0.412	0.055	0.005	0.000	0.000
120	0.018	0.229	0.220	0.454	0.072	0.008	0.000	0.000
160	0.015	0.185	0.240	0.477	0.074	0.008	0.000	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.022	0.083	0.383	0.229	0.249	0.031	0.003	0.000
120	0.021	0.062	0.401	0.242	0.239	0.033	0.001	0.000
160	0.016	0.044	0.421	0.189	0.285	0.044	0.001	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-7: Simulated Log-Likelihood, $\hat{\xi}_{LR}$ sequential test rejection statistics up to 8 FDC, see equation (1.10), for I(1) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.877	0.104	0.017	0.001	0.000	0.000	0.000	0.000
120	0.856	0.122	0.020	0.002	0.000	0.000	0.000	0.000
160	0.877	0.101	0.021	0.002	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.022	0.832	0.125	0.019	0.001	0.000	0.000	0.000
120	0.013	0.833	0.133	0.020	0.001	0.000	0.000	0.000
160	0.010	0.827	0.142	0.019	0.002	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.148	0.120	0.599	0.116	0.016	0.000	0.000	0.000
120	0.100	0.113	0.652	0.120	0.013	0.001	0.000	0.000
160	0.133	0.060	0.669	0.123	0.015	0.001	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.056	0.345	0.237	0.305	0.053	0.003	0.000	0.000
120	0.048	0.215	0.185	0.455	0.087	0.009	0.001	0.000
160	0.043	0.139	0.182	0.523	0.104	0.009	0.000	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.103	0.176	0.352	0.188	0.156	0.023	0.001	0.000
120	0.073	0.089	0.351	0.182	0.259	0.043	0.002	0.000
160	0.055	0.054	0.320	0.187	0.328	0.052	0.003	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-8: Simulated Log-Likelihood, $\hat{\xi}_{LR}$, sequential test rejection statistics up to 8 FDC, see equation (1.10), for I(1) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.862	0.112	0.022	0.004	0.000	0.000	0.000	0.000
120	0.861	0.115	0.022	0.002	0.000	0.000	0.000	0.000
160	0.857	0.120	0.022	0.001	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.019	0.811	0.147	0.022	0.001	0.000	0.000	0.000
120	0.013	0.817	0.150	0.019	0.001	0.000	0.000	0.000
160	0.006	0.818	0.153	0.021	0.002	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.111	0.083	0.662	0.124	0.018	0.001	0.000	0.000
120	0.129	0.082	0.647	0.124	0.017	0.001	0.000	0.000
160	0.105	0.058	0.685	0.134	0.016	0.002	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.072	0.184	0.179	0.465	0.088	0.010	0.001	0.000
120	0.051	0.153	0.150	0.525	0.108	0.011	0.001	0.000
160	0.042	0.154	0.142	0.515	0.131	0.013	0.002	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.080	0.081	0.301	0.161	0.316	0.058	0.004	0.000
120	0.067	0.060	0.241	0.158	0.386	0.078	0.009	0.000
160	0.052	0.043	0.206	0.159	0.455	0.077	0.006	0.001

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-9: Simulated Wald, $\hat{\xi}_W$, sequential test rejection statistics up to 8 FDC, see equation (1.11), for I(1) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.843	0.123	0.030	0.005	0.000	0.000	0.000	0.000
120	0.829	0.132	0.033	0.006	0.000	0.000	0.000	0.000
160	0.861	0.111	0.025	0.003	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.023	0.797	0.147	0.030	0.003	0.000	0.000	0.000
120	0.015	0.804	0.150	0.027	0.003	0.000	0.000	0.000
160	0.012	0.811	0.149	0.024	0.003	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.107	0.123	0.612	0.133	0.022	0.002	0.000	0.000
120	0.053	0.112	0.669	0.143	0.020	0.002	0.000	0.000
160	0.068	0.061	0.692	0.154	0.022	0.002	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.038	0.303	0.251	0.330	0.071	0.006	0.000	0.000
120	0.029	0.190	0.198	0.469	0.100	0.013	0.001	0.000
160	0.019	0.115	0.184	0.541	0.122	0.018	0.001	0.001
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.073	0.178	0.340	0.203	0.175	0.028	0.002	0.000
120	0.042	0.090	0.342	0.201	0.268	0.052	0.004	0.000
160	0.034	0.052	0.304	0.192	0.351	0.062	0.005	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-10: Simulated Wald, $\hat{\xi}_W$, sequential test rejection statistics up to 8 FDC, see equation (1.11), for I(1) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.827	0.127	0.039	0.005	0.001	0.001	0.000	0.000
120	0.833	0.127	0.032	0.007	0.001	0.000	0.000	0.000
160	0.838	0.121	0.034	0.006	0.001	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.021	0.777	0.165	0.031	0.006	0.000	0.000	0.000
120	0.015	0.774	0.175	0.029	0.006	0.001	0.000	0.000
160	0.010	0.791	0.169	0.026	0.004	0.001	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.043	0.079	0.679	0.163	0.029	0.007	0.000	0.000
120	0.052	0.067	0.681	0.161	0.029	0.007	0.002	0.000
160	0.045	0.052	0.709	0.156	0.032	0.005	0.001	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.017	0.134	0.176	0.519	0.126	0.022	0.004	0.001
120	0.012	0.114	0.134	0.563	0.143	0.028	0.006	0.001
160	0.006	0.109	0.132	0.561	0.159	0.027	0.007	0.001
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.030	0.067	0.275	0.178	0.356	0.079	0.013	0.001
120	0.014	0.034	0.216	0.168	0.435	0.114	0.017	0.001
160	0.007	0.031	0.179	0.167	0.484	0.115	0.014	0.003

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-11: Simulated Lagrange Multiplier, $\hat{\xi}_{LM}$, sequential test rejection statistics up to 8 FDC, see equation (1.12), for I(1) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.871	0.105	0.021	0.003	0.000	0.000	0.000	0.000
120	0.849	0.120	0.027	0.004	0.000	0.000	0.000	0.000
160	0.872	0.104	0.021	0.003	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.026	0.822	0.130	0.020	0.002	0.000	0.000	0.000
120	0.015	0.821	0.138	0.024	0.003	0.000	0.000	0.000
160	0.014	0.826	0.139	0.019	0.002	0.000	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.125	0.126	0.618	0.112	0.017	0.001	0.000	0.000
120	0.059	0.123	0.672	0.129	0.016	0.001	0.000	0.000
160	0.072	0.065	0.705	0.139	0.016	0.002	0.000	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.047	0.332	0.254	0.311	0.055	0.003	0.000	0.000
120	0.034	0.209	0.201	0.459	0.086	0.011	0.001	0.000
160	0.020	0.123	0.186	0.542	0.111	0.016	0.001	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.085	0.193	0.351	0.193	0.157	0.021	0.001	0.000
120	0.046	0.097	0.362	0.197	0.250	0.046	0.003	0.000
160	0.038	0.059	0.318	0.191	0.334	0.055	0.005	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-12: Simulated Lagrange Multiplier, $\hat{\xi}_{LM}$, sequential test rejection statistics up to 8 FDC, see equation (1.12), for I(1) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.856	0.113	0.028	0.003	0.001	0.000	0.000	0.000
120	0.851	0.116	0.028	0.004	0.000	0.000	0.000	0.000
160	0.854	0.113	0.028	0.005	0.000	0.000	0.000	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.022	0.814	0.140	0.020	0.003	0.000	0.000	0.000
120	0.016	0.800	0.155	0.024	0.004	0.000	0.000	0.000
160	0.011	0.809	0.156	0.021	0.003	0.001	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.054	0.087	0.700	0.133	0.021	0.004	0.000	0.000
120	0.056	0.068	0.695	0.148	0.026	0.005	0.002	0.000
160	0.049	0.053	0.719	0.147	0.027	0.004	0.001	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.023	0.158	0.181	0.513	0.107	0.015	0.002	0.000
120	0.015	0.130	0.138	0.558	0.131	0.023	0.004	0.001
160	0.007	0.119	0.135	0.563	0.147	0.023	0.006	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.037	0.077	0.312	0.172	0.329	0.065	0.008	0.000
120	0.015	0.041	0.238	0.167	0.423	0.102	0.013	0.001
160	0.010	0.034	0.189	0.168	0.485	0.099	0.013	0.003

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-13: Simulated BIC rejection statistics up to 8 FDC, see Section 7, for I(0) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.560	0.186	0.144	0.075	0.026	0.008	0.001	0.000
120	0.626	0.182	0.119	0.054	0.015	0.004	0.000	0.000
160	0.668	0.167	0.107	0.042	0.013	0.002	0.001	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.013	0.589	0.220	0.120	0.041	0.014	0.002	0.001
120	0.009	0.640	0.205	0.106	0.031	0.008	0.001	0.000
160	0.008	0.668	0.206	0.086	0.027	0.006	0.000	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.029	0.085	0.532	0.241	0.090	0.021	0.002	0.000
120	0.030	0.065	0.593	0.227	0.071	0.013	0.002	0.000
160	0.017	0.055	0.617	0.226	0.069	0.015	0.001	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.008	0.100	0.169	0.467	0.196	0.054	0.007	0.000
120	0.003	0.080	0.170	0.510	0.190	0.042	0.006	0.000
160	0.001	0.089	0.133	0.544	0.192	0.036	0.005	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.013	0.047	0.226	0.248	0.345	0.108	0.014	0.000
120	0.005	0.025	0.164	0.198	0.464	0.128	0.017	0.001
160	0.005	0.016	0.150	0.199	0.486	0.131	0.014	0.000

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-14: Simulated BIC rejection statistics up to 8 FDC, see Section 7, for I(0) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.514	0.236	0.144	0.063	0.031	0.010	0.003	0.000
120	0.570	0.238	0.116	0.049	0.020	0.006	0.001	0.000
160	0.615	0.227	0.096	0.041	0.015	0.005	0.001	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.014	0.539	0.256	0.124	0.048	0.015	0.003	0.000
120	0.010	0.598	0.250	0.094	0.036	0.010	0.002	0.000
160	0.007	0.621	0.245	0.089	0.029	0.008	0.002	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.025	0.081	0.517	0.244	0.098	0.032	0.004	0.000
120	0.039	0.075	0.573	0.220	0.073	0.018	0.002	0.000
160	0.039	0.059	0.597	0.225	0.065	0.013	0.001	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.007	0.105	0.170	0.471	0.184	0.052	0.011	0.001
120	0.003	0.101	0.164	0.495	0.183	0.046	0.006	0.000
160	0.004	0.098	0.181	0.502	0.171	0.040	0.005	0.000
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.004	0.025	0.205	0.223	0.392	0.128	0.022	0.002
120	0.005	0.019	0.246	0.231	0.377	0.105	0.016	0.000
160	0.005	0.019	0.243	0.198	0.400	0.123	0.011	0.001

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-15: Simulated BIC rejection statistics up to 8 FDC, see Section 7, for I(1) series under skip sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.564	0.160	0.172	0.067	0.027	0.009	0.001	0.000
120	0.620	0.164	0.145	0.048	0.018	0.003	0.002	0.000
160	0.676	0.143	0.123	0.038	0.014	0.004	0.001	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.008	0.572	0.231	0.124	0.050	0.013	0.002	0.000
120	0.005	0.635	0.211	0.103	0.037	0.009	0.000	0.000
160	0.004	0.667	0.203	0.084	0.032	0.008	0.002	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.020	0.055	0.551	0.239	0.101	0.030	0.004	0.000
120	0.012	0.063	0.599	0.232	0.073	0.018	0.002	0.000
160	0.018	0.033	0.629	0.222	0.074	0.020	0.002	0.000
	Four Frequency Dependent Coefficients under the null							
80	0.005	0.108	0.188	0.447	0.191	0.054	0.007	0.001
120	0.004	0.066	0.131	0.533	0.205	0.057	0.005	0.000
160	0.002	0.045	0.131	0.564	0.205	0.045	0.006	0.001
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.008	0.040	0.197	0.232	0.372	0.131	0.019	0.001
120	0.006	0.018	0.167	0.187	0.445	0.157	0.018	0.001
160	0.007	0.010	0.159	0.175	0.486	0.144	0.019	0.001

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Skip sampling takes one observation of low frequency variable every three observations of high frequency variable. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Table C.1-16: Simulated BIC rejection statistics up to 8 FDC, see Section 7, for I(1) series under simulated GDP sampling, $s = 3$, and with sample size of {80, 120, 160} observations for the low frequency variable.

	One Frequency Dependent Coefficient under the null							
n	1	2	3	4	5	6	7	8+
80	0.526	0.165	0.181	0.074	0.037	0.014	0.003	0.000
120	0.612	0.163	0.143	0.049	0.025	0.006	0.002	0.000
160	0.642	0.152	0.136	0.044	0.020	0.004	0.001	0.000
	Two Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.007	0.535	0.251	0.133	0.052	0.019	0.002	0.000
120	0.004	0.598	0.230	0.108	0.043	0.013	0.003	0.000
160	0.004	0.642	0.219	0.097	0.030	0.008	0.001	0.000
	Three Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.011	0.033	0.539	0.260	0.112	0.037	0.008	0.000
120	0.017	0.041	0.574	0.240	0.092	0.028	0.007	0.001
160	0.016	0.034	0.619	0.224	0.081	0.020	0.006	0.001
	Four Frequency Dependent Coefficients under the null							
80	0.002	0.047	0.118	0.495	0.234	0.085	0.018	0.001
120	0.003	0.048	0.098	0.542	0.216	0.073	0.016	0.003
160	0.001	0.050	0.105	0.543	0.223	0.060	0.017	0.002
	Five Frequency Dependent Coefficients under the null							
n	1	2	3	4	5	6	7	8+
80	0.004	0.019	0.149	0.170	0.442	0.179	0.036	0.002
120	0.003	0.009	0.114	0.145	0.495	0.194	0.037	0.003
160	0.002	0.009	0.101	0.146	0.537	0.166	0.034	0.004

Notes: Series DGPs are describe by equation (1.20). The frequencies that divide the band spectrum sets are randomly assigned between $[0.10\pi, 0.8\pi]$ with a minimum distance between two consecutive breaks of 0.10π. The values of β_l are randomly assigned on $[-2, 2]$ set with equal probability, but with $|\beta_l - \beta_{l+1}| > 0.5$. For q from 1 to 5. Simulated GDP sampling takes aggregated observations of the low frequency variable, as described in Section 7. Low frequency variable sample size of {80, 120, 160} in a Monthly-Quarterly sample scheme represents, respectively, {20, 30, 40} years of data. With 5,000 MC repetitions.

Chapter 2

Nonparametric short- and long-run Granger-causality testing in the frequency domain

Abstract

This paper proposes a novel nonparametric frequency Granger-causality test. Before testing for causality absence of one series on to another, we apply a filtering process, removing any presence of reverse causality. Then, performing a local kernel regression for each frequency, we can test the hypothesis of non-causality from the distance between each estimate to zero. We provide asymptotic results for strict stationary series respecting α-mixing conditions. Monte Carlo experiments illustrate that our method has good finite sample properties, with overall comparable performance with other alternative methods present in the literature, and superior performance whenever the tested model presents smooth transition coefficients in the frequency domain. Finally, we test causality of term spread and money stock (M2) on real economic growth, as well as, between Monetary Policy Variables and Stock Prices.

Keywords: Granger-causality, Frequency-domain, Nonparametric test, α-mixing.

JEL classification: C12, C14, C22.

1 Introduction

Geweke (1982), Hosoya (1991), and more recently Breitung and Candelon (2006), have proposed methods based on VAR models to measure pointwise causality in the frequency domain. Here, we present a nonparametric alternative to these methods that is not only model-free but it is also robust to a wide range of series dependence structures.

Our method consists of two steps, where the first is performed globally in the time domain and the last locally in the frequency domain. In our first step, we project both the endogenous and the exogenous variable, y_t and x_{t-1}, on past values of the endogenous one, filtering away any feedback causality between those variables. We assume that the number of lags increases as the number of observation increases but slowly. In the second step, we locally regress, frequency by frequency, the filtered series, thus allowing that the relationship between these variables can be different at each frequency.

Our test statistic is flexible in the sense that it is possible to infer causality between any two series, $w_x(\lambda) \to w_y(\lambda)$; conditional causality, $w_x(\lambda) \to w_y(\lambda)|w_z(\lambda)$; and multivariate causality, $\{w_{x_1}(\lambda), \ldots, w_{x_p}(\lambda)\} \to w_y(\lambda)|w_z(\lambda)$. The test statistics for the first and for second hypotheses converge, under the null, at every frequency, to a χ_1^2 distribution, and for latest hypothesis to a χ_p^2 distribution. Furthermore, if parametric models rely on information criteria, such AIC and BIC, to determine the underline model, nonparametric models rely on the selection of appropriate bandwidth values. We suggest the use of the 'leave-one-out' cross-validation method as a data-driven bandwidth selection.

According to Granger (1963), the study of causality in frequency domain provides useful information since economic theory predicts distinct long and short-run relationships between series. Granger also raises the concern that an overall causality test may turn into a misleading result. Granger (1969) further develops the study of the causality measurement via cross-spectrum. In the paper's conclusion, Granger mentions that the results were obtained via parametric modeling and that a direct method to estimate the components of the cross-spectrum 'worth investigating'.

Geweke (1982, 1984) proposed to disentangle the linear feedback into three folds, the causal relationship between the first time series into the second, the inverse causal relation between them and their instantaneous causality. Hosoya (1991), working with second-order stationary processes, also defined three alternative measurements of causality: the measure of association, the one-way measure of association, and the reciprocity measurement. Breitung and Candelon (2006) based on the work of Geweke and Hosoya, proposed a simplified causality test in the frequency domain. The authors circumvent the estimation of non-linear covariance matrix elements of previous methods using a set of trigonometric restrictions. In

Section 2 we present a brief comparison between our approach and these methods. Recently, Breitung and Schreiber (2017) proposed to test the absence of causality for a unique frequency within a frequency band. They also proposed a method for the identification of the phase shift whenever Granger-causality is present. Based on Pierce (1979), Lemmens et al. (2008) proposed a simple Granger-causality test in frequency domain, which has similar power and size to Breitung and Candelon (2006) test.

On the side of the nonparametric Granger-causality literature, Hidalgo (2000, 2005), using the Hannan-inefficient, HI, estimator, proposed a nonparametric test for series that may present long-range dependence, showing that whenever the series exhibit long-memory the HI estimator is more indicated than least squares. Using a kernel-based approach in the frequency domain, Hong et al. (2009) proposed a test to detect extreme downside risk spillover between financial markets, with a standard Normal distribution under the null. Nishiyama et al. (2011) explored non-linear causality and causality in higher conditional moments. Their test can be understood as an omitted-variable test with nontrivial power against $\sqrt{T}$ local alternatives.

One can interpret measurements of causality at different horizons, see Dufour and Taamouti (2010), as equivalent to a frequency domain approaches. An indicative of causality at lag $p = \infty$ implies in causality at frequency 0 and as at lag $p = 1$ to close to frequency π. The main difference between both approaches is that besides the analysis of different horizons is straightforward, the estimation of lags approaching the infinity becomes unreliable. This issue does not impact frequency domain methods, but, the interpretation of causality at a particular frequency might not be clear to all audiences.

Some notations used through the paper: The variable Z_n represents the vector $\{z_t\}_{t=1}^{n}$. Let $w_z(\lambda) = WZ$, where W, $n \times n$, represents the discrete Fourier transform with row j given by $W_j = n^{-1/2}(1, e^{-i\lambda_j}, e^{-i2\lambda_j}, \ldots, e^{-i(n-1)\lambda_j})$, $\lambda_j = 2\pi j/n$, $j = 0, \ldots, n-1$. A spectral kernel function centered at origin and with bandwidth equals to b is represent by $K_b(\lambda) = bK(b\lambda)$, such that if $K(\lambda) = 0$, $|\lambda| > \pi$, then for $b > 1$, $K_b(\lambda) = bK(b\lambda) = 0$ for $|\lambda| > \pi/b$. The prime symbol, as in $w_z(\lambda)'$, means transpose conjugated. Finally, a variable/coefficient with 0 superscript, e.g., z_t^0, represent a variable/coefficient under $\mathbb{H}_0$.

The organization of the article is the following. Section 2 introduces the nonparametric estimation and the test procedures. Section 3 studies the assumptions and the asymptotic theory. Section 4 discusses about estimation procedures. Section 5 presents some considerations about finite sample size properties and Monte Carlo experiments designs. Section 6 reproduces Breitung and Candelon (2006) and Dufour and Tessier (2006) empirical application. Section 7 concludes the paper. Proofs are present in the Appendix 2.A.

2 Nonparametric Granger-causality test

Let us present two representations of the same nonparametric model, one in time domain and other in frequency domain. In time domain, y_t is generated by

$$y_t = \sum_{j=1}^{\infty} \theta_j y_{t-j} + \sum_{j=1}^{\infty} \beta_j x_{t-j} + u_t, \quad E[u_t | x_s, y_s, -\infty < s < t] = 0,$$

where the series $\{x_t, y_t, u_t\}_{t=1}^{\infty}$ are a strict stationary process respecting α-mixing conditions with up to eighth moment bounded and zero mean and β_j is a $1 \times p$ vector. In frequency domain we have that

$$w_y(\lambda) = g_{\vec{yy}}(\lambda) w_{y_{-1}}(\lambda) + g_{\vec{xy}}(\lambda) w_x(\lambda) + u(\lambda), \tag{2.1}$$

where we assume that $g_{\vec{yy}}(\lambda) = \sum_{j=1}^{\infty} \theta_j e^{-i\lambda(j-1)}$ and $g_{\vec{xy}}(\lambda) = \sum_{j=1}^{\infty} \beta_j e^{-i\lambda j}$ converges for all λ. Also, $w_{y_{-1}}(\lambda)$ is the Fourier transform of $\{y_{t-1}\}_{t=-\infty}^{\infty}$.

The coefficient $g_{\vec{zr}}(\lambda)$ refers to the causal effect of past values of z_t on r_t at frequency λ. Thus, the non causality null hypothesis of interest is such that

$$\mathbb{H}_0 : g_{\vec{xy}}(\lambda) = 0, \quad \lambda \in [0, \pi],$$

i.e., the past values of x_t does not Granger-cause y_t series at frequency λ given the past values of y_t, with the alternative hypothesis, $\mathbb{H}_1$, given by $g_{\vec{xy}}(\lambda) \neq 0$, for some $\lambda \in [0, \pi]$.

We can estimate $g_{\vec{xy}}(\lambda)$ by the minimization of the local sum of squares residuals, i.e., given a kernel function, K, and a value of h is such that as n increases nh also increases, but at a lower rate, i.e. $h \to 0$, we have,

$$\min_{\psi_{\vec{yy}}(\lambda), \psi_{\vec{xy}}(\lambda) \in \mathbb{C}} \sum_{\lambda_s \in [-\pi, \pi]} K_{nh}(\lambda - \lambda_s) \left| w_y(\lambda_s) - \psi_{\vec{yy}}(\lambda) w_{y_{-1}}(\lambda) - \psi_{\vec{xy}}(\lambda) w_x(\lambda_s) \right|^2, \tag{2.2}$$

However, as the Fourier transform of past values of y_t was wrote in terms of its first lag frequency domain representation, we can also write $g_{\vec{yy}}(\lambda) w_{y_{-1}}(\lambda)$ as $g_{\vec{yy}}(\lambda) e^{-i\lambda} w_y(\lambda)$. Although this functional dependence between the dependent variable and a regressor might not be problematic in a band spectrum regression for some specification of $g_{\vec{xy}}(\lambda)$, it means that we cannot remove the influence of past values of y_t and consistently estimate $g_{\vec{xy}}(\lambda)$ around a fixed λ. Beyond that, to account for situations where past values of y_t cause x_t, we re-write the exogenous variable as $x_t = x_t^0 + \sum_{j=0}^{\infty} \phi_j y_{t-j}$, or as $w_x(\lambda) = w_x^0(\lambda) + g_{\vec{yx}}(\lambda) w_y(\lambda)$ in the frequency domain, where the elements of RHS are orthogonal to each other. Thus, we have an unfeasible decomposition of (2.1) RHS as

$$w_y(\lambda) = g_{\vec{yy}}(\lambda) w_{y_{-1}}(\lambda) + \left[g_{\vec{xy}}(\lambda) w_x^0(\lambda) + g_{\vec{xy}}(\lambda) g_{\vec{yx}}(\lambda) w_y(\lambda) \right] + u(\lambda). \tag{2.3}$$

Nevertheless, to consistently estimate $g_{\vec{xy}}(\lambda)$ in (2.3) at each λ, we project x_{t-1} and y_t on the span of past values of y_t before performing the local regression. This step performs a time-domain global filtering on (2.3), removing past values of y_t influence. Thus, the unfeasible model (2.3) becomes,

$$w_y^0(\lambda) = g_{\vec{xy}}(\lambda) w_x^0(\lambda) + w_u(\lambda), \tag{2.4}$$

where $y_t = y_t^0 + \sum_{j=1}^{\infty} \theta_j^0 y_{t-j}$ and the two terms on RHS are orthogonal to each other. Now, we can test locally the absence of causality, between $w_x^0(\lambda)$ and $w_y^0(\lambda)$, without the interference of other terms in the frequency domain. Summarizing, for $\{X, Y\} = \{x_{t-1}, y_t\}_{t=q+1}^n$, the proposed two-step procedure is the following:

Step 1: Let Y_q be a, $n - q \times q$, matrix of past values of y_t, such that row j, $j \in [q+1, n]$, is defined as $Y_{q,j} = [y_{j-1}, \ldots, y_{j-q}]$, and the projection matrix $P_Y = Y_q (Y_q' Y_q)^{-1} \, Y_q'$. Define $\hat{X}^0 = (1 - P_Y) X$ and $\hat{Y}^0 = (1 - P_Y) Y$.

Step 2: Nonparametrically estimate $g_{\vec{xy}}(\lambda)$ and test $\mathbb{H}_0$ by measuring the distance from this estimate to zero, for $\lambda \in [0, \pi]$, using the approximated model (2.4).

Therefore, for $\hat{X}^0$, $n \times p$, and $\hat{Y}^0$, $n \times 1$, we have,

$$\min_{\phi_{\vec{xy}}(\lambda) \in \mathbb{C}} \sum_{\lambda_s \in [-\pi, \pi]} K_{\tilde{n}h}(\lambda - \lambda_s) \left| w_{\hat{y}}^0(\lambda_s) - \phi_{\vec{xy}}(\lambda)' w_{\hat{x}}^0(\lambda_s) \right|^2,$$

with a feasible solution for every λ given by,

$$\hat{g}_{\vec{xy}}(\lambda) \;=\; \hat{f}_{\hat{x}^0 \hat{x}^0}^{-1}(\lambda) \hat{f}_{\hat{x}^0 \hat{y}^0}(\lambda), \tag{2.5}$$

where

$$\hat{f}_{zr}(\lambda) \;=\; \frac{1}{\tilde{n}} \sum_{\lambda_s \in [-\pi, \pi]} K_{\tilde{n}h}(\lambda - \lambda_s) w_r(\lambda_s) w_z(\lambda_s)', \tag{2.6}$$

are smoothed spectral estimates and $\tilde{n} = n - q$.

The asymptotic variance matrix estimate of $g_{\vec{xy}}(\lambda)$, $p \times p$, is defined under $\mathbb{H}_0$ as

$$\hat{\Sigma}_{\vec{xy}}(\lambda) = \hat{f}_{\hat{x}^0 \hat{x}^0}^{-1}(\lambda) \hat{f}_{\hat{u}^0 \hat{u}^0}(\lambda) \int_{-\infty}^{\infty} k(s)^2 ds, \tag{2.7}$$

where $w_{\hat{u}}(\lambda) = w_{\hat{y}}^0(\lambda) - \hat{g}_{\vec{xy}}(\lambda) w_{\hat{x}}^0(\lambda_s)$. Thus, the null hypothesis that the past values of x do not cause y at frequency λ can be inferred by the test statistic

$$W_\lambda = \hat{g}_{\vec{xy}}(\lambda)' \left(\hat{\Sigma}_{\vec{xy}}(\lambda) h \right)^{-1} \hat{g}_{\vec{xy}}(\lambda), \quad \lambda \in [0, \pi],$$

or for each $i = 1, \ldots, p$,

$$W_{i,\lambda} = \frac{\hat{g}_{i,\overrightarrow{xy}}(\lambda)^2}{\left(\hat{\Sigma}_{ii,\overrightarrow{xy}}(\lambda)h\right)}, \quad \lambda \in [0, \pi].$$

Allowing q to diverge in the first step, we indirectly estimate $g_{\overrightarrow{yy}}(\lambda)$ and $g_{\overrightarrow{yx}}(\lambda)$ by the infinite summation of coefficients θ_j^0 and α_j weighted by $e^{-ij\lambda}$. Thus, in (2.5) we have a measurement of the pointwise causality of past values of x_t on y_t for a given λ.

In comparison, the causality tests of Geweke (1982) and Hosoya (1991) are based on a parametric representation with a finite number of lags, say l. Let $z_t = [y_t, x_t]'$ and $\Theta(L)z_t = \varepsilon_t$, the VMA representation is given by $z_t = \Phi(L)\varepsilon_t$. Now, let the Cholesky decomposition of the inverse covariance matrix, $\Sigma = E(\varepsilon_t\varepsilon_t')$, be given by $\Sigma^{-1} = G'G$, where G is a lower triangular matrix, $\Psi(L) = \Phi(L)G^{-1}$, and $\eta_t = G\varepsilon_t$ then

$$z_t = \Psi(L)\eta_t = \begin{bmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ \Psi_{21}(L) & \Psi_{22}(L) \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}.$$

The measurement of causality is given by

$$M_{x\to y}(\lambda) = \log\left[\frac{f_{yy}(\lambda)}{f_{yy|\eta_1}(\lambda)}\right] = \log\left[1 + \frac{|\Psi_{12}(e^{-i\lambda})|^2}{|\Psi_{11}(e^{-i\lambda})|^2}\right],$$

where $f_{yy|\eta_1}(\lambda)$ is the spectrum of y_t given η_{1t}. Past values of x_t does not cause y_t at frequency λ if $|\Psi_{12}(e^{-i\lambda})| = 0$. Breitung and Candelon (2006) measurement of causality is based on the same principle, but using a set of linear restrictions rather than estimating $\Psi(L)$ matrix. For the finite lags linear model described by

$$y_t = \alpha_1 y_{t-1} + \cdots + \alpha_l y_{t-l} + \beta_1 x_{t-1} + \cdots + \beta_l x_{t-l} + \varepsilon_{1t},$$

Breitung and Candelon (2006) define the null hypothesis that the past values of x_t does not cause y_t at frequency λ as $\mathbb{H}_0 : R(\lambda)\beta = 0$, where

$$R(\lambda) = \begin{bmatrix} \cos(\lambda) & \cos(2\lambda) & \ldots & \cos(l\lambda) \\ \sin(\lambda) & \sin(2\lambda) & \ldots & \sin(l\lambda) \end{bmatrix}.$$

Notice that the approach of Breitung and Candelon (2006) can be understood as an approximation of $g_{\overrightarrow{xy}}(\lambda)$ by a finite sum of l elements rather than our local approach. Thus, we can state that despite our nonparametric test and the parametric tests above mentioned exploit the same type of information, our method allows for dependence structures with no upper bound on the number of lags involved and provides valid inference rules under this framework. Methodologically, our test measures whether the cross-spectrum between filtered X_n and Y_n are distinct from zero and the previous investigated parametric tests whether the marginal spectra of Y_n and of Y_n given a vector of explanatory variables are different.

Additionally, since we are estimating a vector of coefficients and the covariance matrix, we can also test for multivariate causality. Ladroue et al. (2009) and Barrett et al. (2010) presented cases where the multivariate, or group, analysis of causality has more appeal than the univariate causality measurement, e.g., binding interactions of proteins during the yeast-cell cycle and interactions between groups of neurons using fMRI, functional magnetic resonance imaging, data.

3 Asymptotic Theory

To investigate the asymptotic properties of our causality tests, let us introduce the following conditions:

Assumption 3. *Let Z_n be a vector strictly stationary strongly mixing process with zero mean. The lth order cumulants, $cum(Z_j^{(a_0)}, Z_{j+s_1}^{(a_1)}, \dots, Z_{j+s_{l-1}}^{(a_{l-1})}) = r_{s_1,\dots,s_{k-1}}^{(a_0,a_1,\dots,a_{l-1})}$, are well-defined and for $l = \{1, \dots, 8\}$, $\sum_{s_1,\dots,s_{l-1}} |r_{s_1,\dots,s_{l-1}}^{(a_0,a_1,\dots,a_{l-1})}| < \infty$ holds.*

Assumption 4. *Let $x_t = x_t^0 + \sum_{j=1}^{\infty} \phi_j y_{t-j}$ and $y_t = y_t^0 + \sum_{j=1}^{\infty} \theta_j y_{t-j}$ where $Cov(x_t^0, y_{t-j}) = 0$ and $Cov(y_t^0, y_{t-j}) = 0$, for $j > 0$. And for $q = c_\delta n^\delta$, such that $0 < c_\delta < \infty$ and $0 < \delta < 1/3$, we have $n^{1/2} \sum_{j>q} |\phi_j| \to 0$ and $n^{1/2} \sum_{j>q} |\theta_j| \to 0$ as $n \to \infty$.*

Assumption 5. *Let $f_{xx}(\lambda)$, $f_{xu}(\lambda)$, and $f_{uu}(\lambda)$ be twice continuously differentiable in any open set of $\lambda \in [-\pi, \pi]$. Also, let $f_{x^0x^0}(\lambda)$ be positive definite, and $f_{uu}(\lambda)$ be larger than zero for all λ.*

Assumption 6. *Let k be even, bounded, continuous, $k(0) = 1$, k satisfies $|k(x)| \leq \bar{k}(x)$ where*

$$\int_0^\infty (1 + x^2)\bar{k}(x)dx < \infty$$

and $\bar{k}(x)$ is monotonically decreasing on $[0, \infty)$ and chosen to be even. Finally,

$$K(\lambda) = \frac{1}{2\pi} \int_{\mathbb{R}} k(x)e^{-ix\lambda}dx.$$

Assumption 7. *The bandwidth $h = c_\gamma n^{-\gamma}$, such that $0 < c_\gamma < \infty$, $0 < \gamma < 4/5$ and $\gamma + \delta < 1$.*

Let us briefly discuss our assumptions before introduce some asymptotic results. Before Rosenblatt (1984), other authors, e.g. Hannan (1970), assumed more restrictive conditions under the processes, e.g. iid innovations, on spectral estimation. Rosenblatt extends their results to a larger class of spectral estimates assuming stationary strong mixing process. Assumptions 3 and 5 guarantee the consistency and asymptotic normality of spectrum and the cross-spectrum estimation. According to Bühlmann (1996) a sufficient condition for summability of the cumulants in Assumption 3 up to order eight is given by

$$\sum_{i=0}^{\infty}(i+1)^6\alpha^{\tau/(14+\tau)}(i) < \infty, \quad E||Z_n||^{8+\tau} < \infty, \quad \tau > 0.$$

where α represent the mixing coefficients.

Assumption 4 is also an assumption of Gonçalves and Kilian (2007), which extends the results of Lewis and Reinsel (1985) to an infinite autoregressive α-mixed series. The results of Lütkepohl and Poskitt (1996) for a VAR(∞) Granger-causality test are also based on Lewis and Reinsel (1985) findings, thus the upper bound for the number of lags at order $n^{1/3}$. Assumption 5 guaranties the usual nonparametric estimation procedure needs that the function is twice differentiable, see Robinson (1989), as well the no-multicollinearity condition. Assumption 3 and 5 satisfy Assumption 1 of Gonçalves and Kilian (2007), thus DGPs such as an invertible VARMA(p,q) and VAR(∞) model satisfy these assumptions. Assumptions 6 and 7 satisfy the kernel spectral estimation conditions of Rosenblatt (1984). Robinson (1983) shows that for a kernel k respecting Assumption 6 we have $\int_{-\pi}^{\pi} K(\lambda)d\lambda = 1$ and since k is even, a vast group of spectral window K is satisfied, e.g. Epanechnikov, Uniform, among others. One could employ high order spectral kernels, or top flat lag windows, to obtain a less restrictive condition, $\gamma < 4/5$, on Assumption 7, see Jones and Signorini (1997) and Politis and Romano (1999).

Theorem 6 states the asymptotic distribution of the two-step nonparametric estimation of $\hat{g}_{\vec{xy}}(\lambda)$ for $\lambda \in [0, \pi]$ under $\mathbb{H}_0$.

Theorem 6. *Let (y_t, x_t) satisfy Assumptions 3 to 5. Also, assume after the first step, that the estimation, defined in equation (2.5) on nonparametric model (2.4) respects Assumptions 6 and 7. Then, as $n \to \infty$ under $\mathbb{H}_0$, the asymptotic distribution of the two-step nonparametric estimation of $\hat{g}_{\vec{xy}}(\lambda)$, for $\lambda \in [0, \pi]$, is given by*

$$\sqrt{h^{-1}}\hat{g}_{\vec{xy}}(\lambda) \xrightarrow{d} N(0, \Sigma_{\vec{xy}}(\lambda)), \quad \Sigma_{\vec{xy}}(\lambda) = \eta(\lambda) f^{-1}_{x^0x^0}(\lambda) f_{uu}(\lambda) \int_{-\infty}^{\infty} K(s)^2 ds < \infty,$$

with $\eta(\lambda) = 2$ for $\lambda = j\pi$, $j \in \mathbb{Z}$ and $\eta(\lambda) = 1$ otherwise. The estimator given by *equation (2.7) is consistent for the covariance matrix* $\Sigma_{\overrightarrow{xy}}(\lambda)$*, and*

$$W_\lambda \xrightarrow{p} \chi^2_p, \quad W_{i,\lambda} \xrightarrow{p} \chi^2_1, \quad i = 1, \ldots, p.$$

Theorem 7 states that, given our two-step procedure, the estimates under $\mathbb{H}_1$ are consistent and power tends to one as $n \to \infty$.

Theorem 7. *Let* (y_t, x_t) *satisfy Assumptions 3 to 5. Also, assume after the first step, that the estimation, defined in equation (2.5) on nonparametric model (2.4) respects Assumptions 6 and 7. Then, as* $n \to \infty$ *under* $\mathbb{H}_1$*, the asymptotic distribution of the two-step nonparametric estimation of* $\hat{g}_{\overrightarrow{xy}}(\lambda)$*, for* $\lambda \in [0, \pi]$*, is given by*

$$\sqrt{h^{-1}}(\hat{g}_{\overrightarrow{xy}}(\lambda) - g_{\overrightarrow{xy}}(\lambda)) \xrightarrow{d} N(0, \Sigma_{\overrightarrow{xy}}(\lambda)),$$

where $\Sigma_{\overrightarrow{xy}}$ *is defined in Theorem 6, then*

$$hW_\lambda \xrightarrow{p} c, \quad hW_{i,\lambda} \xrightarrow{p} c_i, \quad i = 1, \ldots, p.$$

for $0 < c, c_i < \infty$*.*

4 Empirical procedures

In theory K is unlimited in λ and concentrated at zero. As n increases $K(0)$ also increases and by Assumption 6 we know that kernel tails goes fast to zero. Thus, for finite sample rather than estimate (2.6) for all spectrum, we focus on a neighborhood $\mathscr{B}_{\tilde{n}h}$ of λ, such that $\lambda_s \in \mathscr{B}_{nh}$ whenever $s \in [j - \lfloor nh \rfloor, j + \lfloor nh \rfloor]$, $\lfloor z \rfloor$ stands for the lower integer close to z, and $\tilde{n} = n - q$.

A pending issue to examine is how the first step influences the measurement of the Granger-causality in finite samples. A higher value of δ certainly will cover the removal of any presence of y_t lags on x_{t-1}. However, as δ increases $\tilde{n}$ decreases, which will result in an efficiency loss. On other hand, γ must be large enough to satisfy Assumption 7 and to control bias in the spectral estimation, which results in a small $\tilde{n}h$. Notwithstanding, if $\tilde{n}h$ is too small the estimation of $\hat{g}_{\overrightarrow{xy}}(\lambda)$ is also compromised. Defining a balance of the pair $\{q, \tilde{n}h\}$ is not trivial. We, here, use $q = c_\delta n^\delta = [2n^{1/3}]$ for empirical and Monte Carlo results, see Lütkepohl and Poskitt (1996).

Robinson (1989) suggest the 'leave-one-out' cross-validation method for chose the optimum bandwidth. After the first step, one can search for h^*, where $h^* =$

$\arg\min_h CV(h)$, $CV = \sum_j [w^0_{\tilde{y}}(\lambda) - \tilde{g}_{\vec{xy}}(\lambda) w^0_{\tilde{x}}(\lambda)]^2$, $\lambda = 2\pi j/n$, and let $\tilde{f}_{zr}(\lambda)$ be equivalent to (2.6) but with λ_s restrict to $\mathscr{B}^-_{\tilde{n}h}$, then

$$\tilde{g}_{\vec{xy}}(\lambda) = \tilde{f}^{-1}_{x^0x^0}(\lambda)\tilde{f}_{x^0y^{0|1}}(\lambda)$$

where $\mathscr{B}^-_{\tilde{n}h}$ denotes the set $\mathscr{B}_{\tilde{n}h}$ with the exclusion of frequency λ. According to Wong (1983), this method measures the average prediction power of $\tilde{g}_{\vec{xy}}(\lambda)$ on the 'new' sample of $w^0_y(\lambda)$. Thus, h^* provides the best forecast given the set of possible values of h. It is also important to note that the choice of the kernel function has less impact in the outcome than the correct specification of the bandwidth, h. The usual kernel choices are the Uniform kernel, $k(v) = 0.5\mathbb{1}_{\{|v|\leq 1\}}$, the Epanechnikov kernel, $k(v) = 0.75(1 - v^2)\mathbb{1}_{\{|v|\leq 1\}}$, and the Gaussian kernel, $k(v) = (\sqrt{2\pi})^{-1}\exp(-v^2/2)$. Other kernel functions of interest are the Triangular, $k(v) = (1-|v|)\mathbb{1}_{\{|v|\leq 1\}}$ and the Cosine, $k(v) = (\pi/4)\cos(v\pi/2)\,\mathbb{1}_{\{|v|\leq 1\}}$. Our Monte Carlo and empirical results are based on Epanechnikov kernel.

Based on the Beltrão and Bloomfield (1987), Robinson (1991b) suggest an alternative cross-validation technique. Given a bandwidth h and the estimated errors before first step, we can compute a 'leave-one-out' sample spectrum of $u(\lambda)$ as

$$\hat{f}^-_{uu}(\lambda) = \frac{2\pi}{n}\sum_{\lambda_s\in\mathscr{B}^-_\lambda} K_{\tilde{n}h}(\lambda - \lambda_s) I_{uu}(\lambda_s), \quad I_{uu}(\lambda_s) = \hat{u}(\lambda_s)'\hat{u}(\lambda_s)$$

and the optimum bandwidth is given by the minimization of

$$L(\hat{f}^-_{uu}) = \int_0^\pi \left(\log \hat{f}^-_{uu}(\lambda) + I_{uu}(\lambda)/\hat{f}^-_{uu}(\lambda)\right) d\lambda$$

Robinson (1991b) showed, under regular conditions as u_t being iid, that minimizing $L(\hat{f}^-_{uu})$ is equivalent to minimize the Integrated Mean Squared Error, IMSE. Both cross-validation techniques, here presented, perform similarly. Results reported on our Empirical and Monte Carlo sections are obtained under the first method. Another alternative is to choose locally the optimum bandwidth for each of spectral estimates, adding more flexibility, as in Velasco (2002).

Another point of concern is the spectrum behavior at frequency 0 since this frequency represents the sample mean for the sample periodogram. We suggest to use $|z(\lambda_1)|$ in place of $z(\lambda_0)$, where $\lambda_j = 2\pi j/n$.

5 Simulations results

To study the performance of size and power of our test, we propose four sets of DGPs. In the first, we aim to investigate the rejection behavior when x_{t-1}

does not causes y_t. In the second set of DGPs, we present two models where the relationships between x_{t-1} and y_t are driven, in the frequency domain, by trigonometric functions. Under this framework, it is possible to observe how the test behaves in the presence of smooth transitions on $g_{\overrightarrow{xy}}(\lambda)$. In the third, we have a set of 3 variable where z_{t-1} and x_{t-1} cause y_t but the former does not cause y_t given x_{t-1}. Moreover, we reproduce, with the fourth DGP set, the Monte Carlo simulation of Breitung and Candelon (2006). In their study, they propose to investigate series where x_{t-1} causes y_t only at some frequencies and not in others. In all studied cases, we compare our results with Breitung and Candelon (2006), henceforth B&C, test.

The first set of DGPs comprehends DGP 1.1 to 1.4. Notice that, some DGPs are not in consonance with B&C test assumptions, a finite VAR model with iid innovations. DGP 1.2 and 1.3 are fitted only with VAR(∞), and DGP 1.4 presents conditional heteroskedasticity. The results of B&C test under the last DGP can be improved using a HAC covariance matrix estimator. B&C test has the limitation that for the fitted VAR(p), p must be greater than 2. In cases where AIC selects $p \leq 2$, we use $p = 3$. Due to the nonparametric nature of our test, it is not restricted to lag based models, neither it suffers from model misspecification of omitted lags. Additionally, in the Section 3, we have showed that our test is consistent to series respecting α-mixing condition, like conditional heteroskedasticity models.

$$\begin{aligned}
\text{DGP 1.1:} \quad & y_t = v_{1t}, \\
& x_t = v_{2t}, \\[1em]
\text{DGP 1.2:} \quad & y_t = v_{1t} - 0.4v_{1t-1}, \\
& x_t = 0.5y_{t-1} + 0.3y_{t-2} + v_{2t} - 0.3v_{2t-1}, \\[1em]
\text{DGP 1.3:} \quad & y_t = 0.6y_{t-1} + v_{1t} - 0.4v_{1t-1}, \\
& x_t = 0.4x_{t-1} + 0.5y_{t-1} + 0.3y_{t-2} + v_{2t} - 0.3v_{2t-1}, \\[2em]
\text{DGP 1.4:} \quad & y_t = 0.6y_{t-1} + u_{1t}, \\
& x_t = 0.4x_{t-1} + 0.5y_{t-1} + 0.3y_{t-2} + u_{2t},
\end{aligned}$$

with $u_t = \sigma_t z_t$, $\sigma_t^2 = 0.01 + 0.2\sigma_{t-1}^2 + 0.7u_{t-1}^2$, $v_t \sim W.N.(0, V)$ and

$$V = \begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix}, \qquad \text{where } \nu = [0.50, 1, 2].$$

Results for different values of ν does not affect the rejection rate significantly. Therefore, we only report results with $\nu = 1$. We present the results for the first DGP set in two fashions. In Figure 2-1 and 2-2, we report visually the

Table 2-1: Average of simulated rejection over all frequencies for DGP 1.1-4.

	DGP 1.1		DGP 1.2		DGP 1.3		DGP 1.4	
n	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C
100	0.097	0.055	0.107	0.056	0.102	0.055	0.124	0.068
250	0.062	0.055	0.068	0.058	0.063	0.057	0.081	0.071
500	0.048	0.057	0.050	0.056	0.050	0.058	0.063	0.070

Notes: $\hat{W}_\lambda$ stands for our nonparametric test and B&C for the Breitung and Candelon (2006) test. Sample size of {100, 250, 500} observations, $\nu = 1$, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

average rejection over the frequencies. B&C test has an overall better size for few observations and frequencies close to the boundaries. Stationary series following what Granger (1966b) called 'the typical shape of an economic variable', does not have a favorable signal-to-noise ratio for high frequencies. The signal concentration in low frequencies results in a negligible amount of information in high frequencies. Thus, spectral methods based on regression techniques will encounter difficulties in those areas. Since DGP.1 has a flat spectrum, the rejection rate is practically constant for all frequencies. The small higher rejection rate close to the boundaries of DGP.1 is a result of the mirroring procedure. B&C test performs relatively well in our tests even some with DGPs not respect its assumptions. Table 2-1 shows the average rejection rate over all frequencies. Despite the uneven concentration of the spectrum, for all DGPs the rejection rate for our nonparametric test seems to converge quickly to the asymptotic significance rate of 5%.

In the second group of DGPs, we study the behavior of our test and B&C test in the face of smooth transitions in the relation between $w_y(\lambda)$ and $w_x(\lambda)$. In DGP 2.1 we have a positive relationship that vanishes as λ approaches π. For DGP 2.2 the relationship changes signal at $\pi/2$, with past values of x_t causing y_t for all frequencies except $\pi/2$.

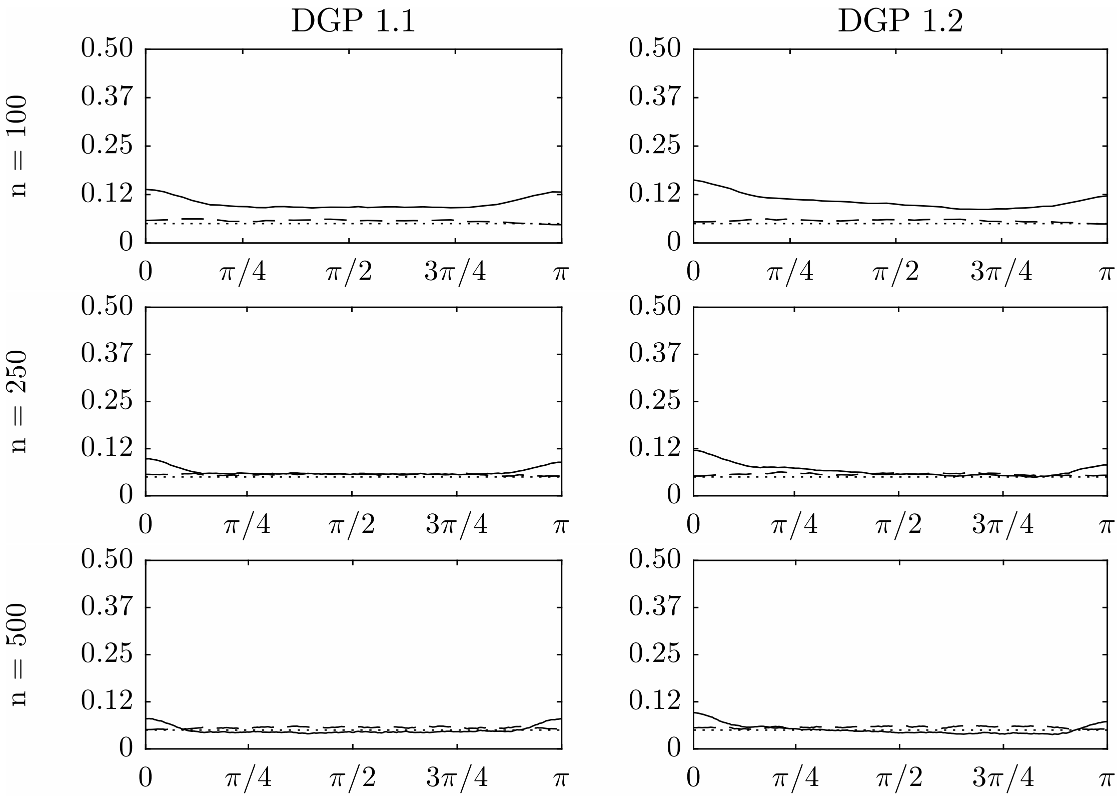
$$\text{DGP 2.1:} \quad w_y(\lambda) = g(\lambda)w_x(\lambda) + w_v(\lambda),\ g(\lambda) = 0.5(1 + \cos(\lambda))$$

$$\text{DGP 2.2:} \quad w_y(\lambda) = g(\lambda)w_x(\lambda) + w_v(\lambda),\ g(\lambda) = \cos(\lambda)$$

with $x_t, v_t \sim W.N.(0, 1)$.

Figure 2-3 reports power results for second group of DGPs. According to the results, our nonparametric model reports an overall higher rejection rate for both models as well as a smooth rejection behavior over the frequencies. Results indicated that the parametric model used by B&C framework does not provide a

Figure 2-1: Simulated size for DGP 1.1 and 1.2.



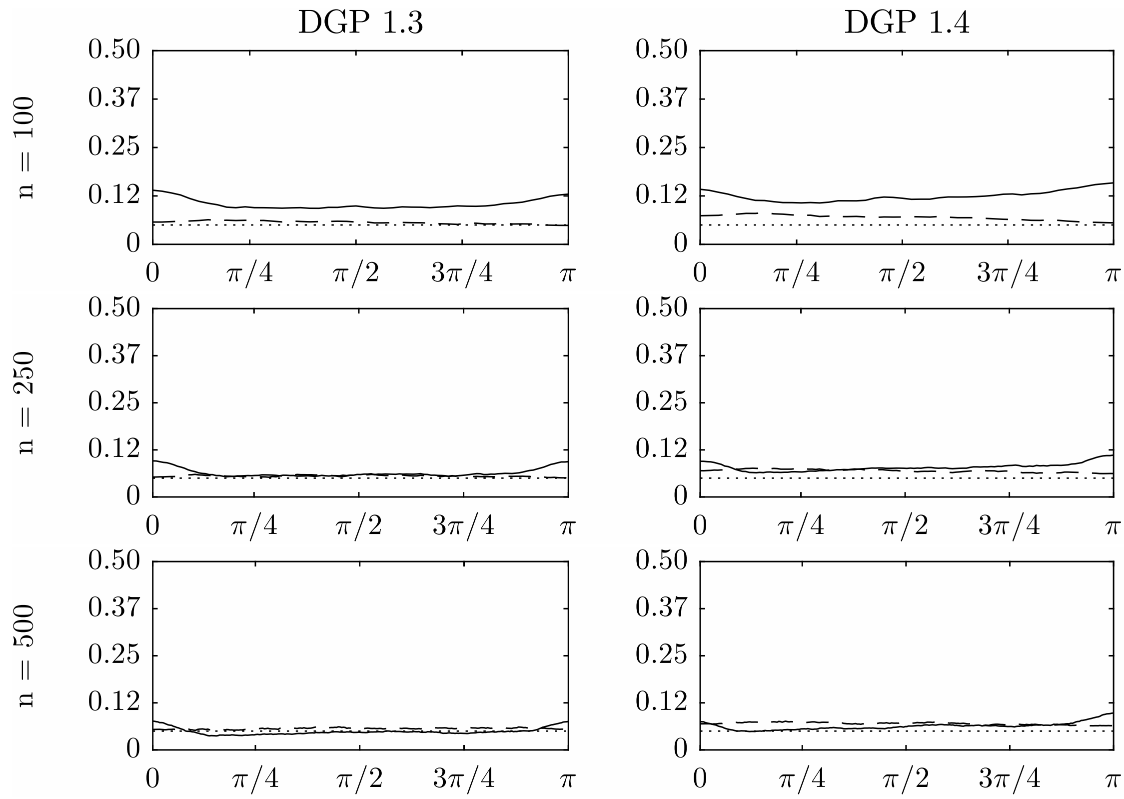
Notes: Solid, dashed, and dash-dot line represent the average rejection rate for our nonparametric test, W_λ, for B&C test, and for our nonparametric test without the Step 1, respectively. Dotted line represent the 5% rejection line. Sample size of {100, 250, 500} observations, $\nu = 1$, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

good approximation of the $g(\lambda)$ function. Contrarily, our nonparametric test is more indicated for this cases since the smooth $g(\lambda)$ function is well approximated by the local weighted regression method.

For the third DGPs sets, DGP 3.1-4, we study the measurement of conditional causality. Given the DGPs' formulation, we expect a rejection rate for x_t close to unity and close to the nominal, 5%, for z_t, for all frequencies.

$$\text{DGP 3.1:} \quad x_t = v_{2t},$$

Figure 2-2: Simulated size for DGP 1.3 and 1.4.



Notes: Solid, dashed, and dash-dot line represent the average rejection rate for our nonparametric test, W_λ, for B&C test, and for our nonparametric test without the Step 1, respectively. Dotted line represent the 5% rejection line. Sample size of {100, 250, 500} observations, $\nu = 1$, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

$$\text{DGP 3.2:} \quad x_t = v_{2t} - 0.3v_{2t-1},$$

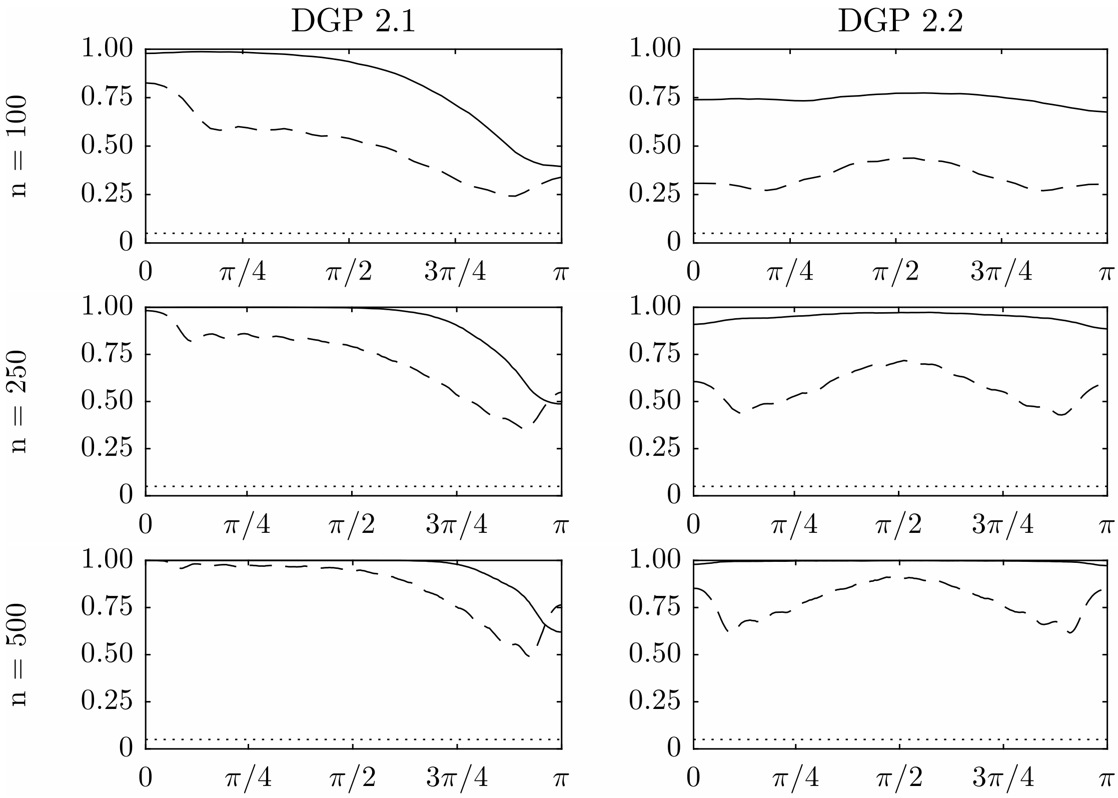
$$\text{DGP 3.3:} \quad x_t = 0.4x_{t-1} + v_{2t} - 0.3v_{2t-1},$$

$$\text{DGP 3.4:} \quad x_t = 0.4x_{t-1} + u_t,$$

with $y_t = 0.5x_{t-1} - 0.3x_{t-2} + v_{1t} - 0.4v_{1t-1}$, $z_t = 0.3x_t + e_t$, $e_t \sim N(0,1)$, $v \sim i.i.d.W.N.(0,1)$, and $u_t = \sigma_t z_t$, $\sigma_t^2 = 0.1 + 0.2\sigma_{t-1}^2 + 0.7u_{t-1}^2$.

Hosoya (2001) suggests projecting the two series of interest in past values of a third one. Thus, the causality of x_{t-1} on y_t given z_t can be measured testing for $\{w_t, m_t\}$, where $w_t = (1 - P_z)x_t$ and $m_t = (1 - P_z)y_t$ and P_z is the projection matrix of past values of z_t. Geweke (1984) suggest to include the information of the third variable in an augment VAR model. Under our framework, the estimation

Figure 2-3: Simulated power for DGP 2.1 and 2.2.



Notes: Solid line and dashed line represent the average rejection rate for our nonparametric test, W_λ, and for B&C test, respectively. Dotted line represent the 5% rejection line. Sample size of $\{100, 250, 500\}$ observations, $\nu = 1$, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

of a multivariate x_t is close to Geweke (1984) suggestion. Table 2-2 presents the result of our test and B&C test using Geweke's procedure. Our test shows a higher rejection rate for both, x_t and z_t, series when compared with B&C test. For small sample sizes, our test reports almost the double rejection rate, when $x_{t-1} \to y_t | z_{t-1}$. In respect to $z_{t-1} \to y_t | x_{t-1}$, our test presents higher rejection, but as n increases the rejection rate approximates to the nominal rate. In fact, experiments with larger values of n show that both tests report similar results.

Finally, the fourth set of DGPs replicates the Breitung and Candelon (2006) Monte Carlo experiment, where the model was constructed around a Gegenbauer polynomial, $b_\omega(L) = [1 - 2cos(\omega)L + L^2]$. This polynomial has as characteristic of no gain at frequency ω, i.e., if $y_t = b_\omega(L)x_{t-1} + u_t$ then x_{t-1} does not cause y_t at frequency ω. DGP 4 exemplify their model for ω equals to $\{\pi/8, \pi/4, 3\pi/8, \pi/2\}$.

Table 2-2: Average of simulated rejection statistics over all frequencies for DGP 3.1-4.

n	DGP 3.1		DGP 3.2		DGP 3.3		DGP 3.4	
	$x_{t-1} \not\to y_t \mid z_{t-1}$							
n	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C
100	0.899	0.543	0.944	0.554	0.891	0.528	0.730	0.375
250	0.978	0.818	0.994	0.807	0.976	0.819	0.909	0.684
500	0.995	0.942	0.999	0.923	0.994	0.946	0.974	0.867
	$z_{t-1} \not\to y_t \mid x_{t-1}$							
n	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C	$\hat{W}_\lambda$	B&C
100	0.136	0.058	0.139	0.060	0.139	0.059	0.129	0.058
250	0.093	0.060	0.096	0.062	0.091	0.062	0.088	0.060
500	0.073	0.063	0.079	0.062	0.072	0.062	0.073	0.065

Notes: $\hat{W}_\lambda$ stands for our nonparametric test and B&C for the Breitung and Candelon (2006) test. Sample size of {100, 250, 500} observations, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

$$\begin{aligned} \text{DGP 4.}\omega\text{:} \quad y_t &= 0.1y_{t-1} + 0.3b_\omega(L)x_{t-1} + \varepsilon_{1t}, \\ x_t &= -y_{t-1} + 0.1x_{t-1} - 0.2x_{t-2} + 0.3x_{t-3} + \varepsilon_{2t}, \end{aligned}$$

where

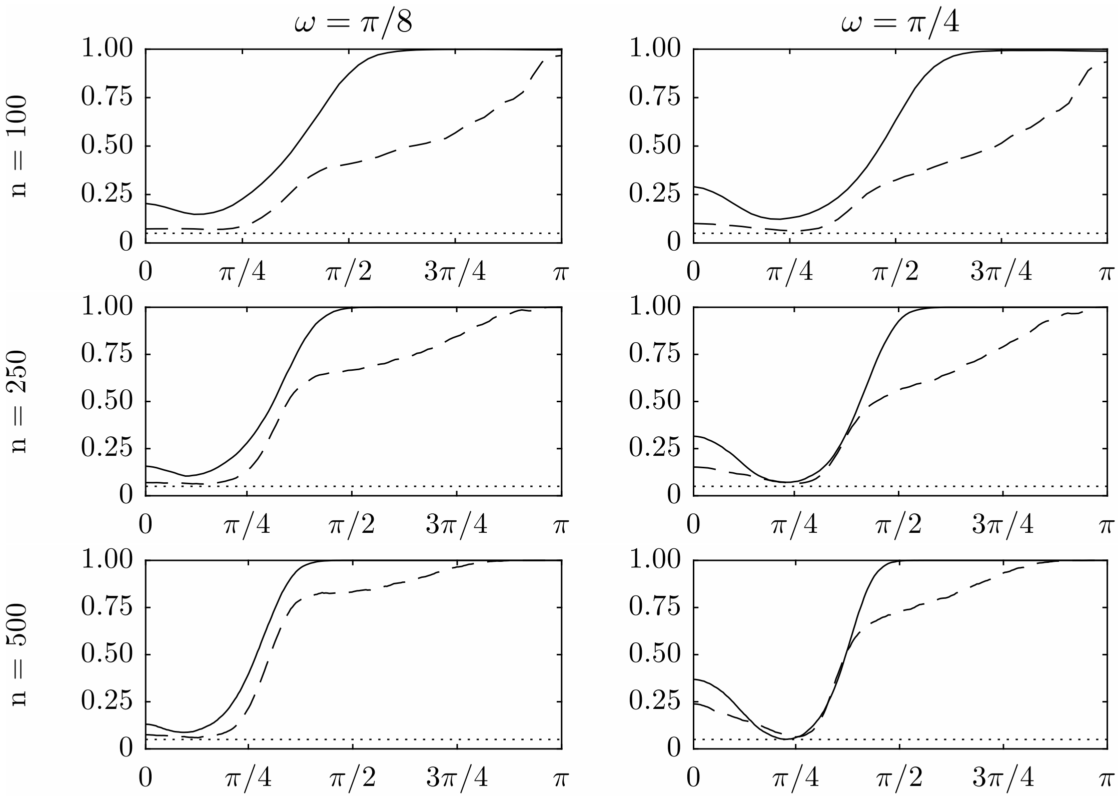
$$\varepsilon_t \sim N(0, V), \quad V = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.5 \end{bmatrix}$$

Breitung and Candelon (2006) based their results on perfect knowledge of the model. Here we use AIC to select the number of lags, with maximum of lags equal to $q = [2n^{1/3}]$ lags. Our test, Figure 2-4 and 2-5, compared with B&C shows, in general, presents better performance over the spectrum. In contrary with the results presented in Breitung and Candelon (2006), the B&C model, based on AIC model selection, reports lower rejection rates for frequencies distinct of ω than our nonparametric model.

6 Empirical results

Here, we revisit Breitung and Candelon (2006) and Dufour and Tessier (2006) empirical applications. The first assess the causality of interest rate spreads on real economic growth. The latter, using a set of macroeconomic policy variables

Figure 2-4: Simulated power for DGP 4.ω with $\omega = \{\pi/8, \pi/4\}$.



Notes: Solid line and dashed line represent the average rejection rate for our nonparametric test, $\hat{W}_\lambda$, and for B&C test, respectively. Dotted line represent the 5% rejection line. Sample size of $\{100, 250, 500\}$ observations, $\nu = 1$, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

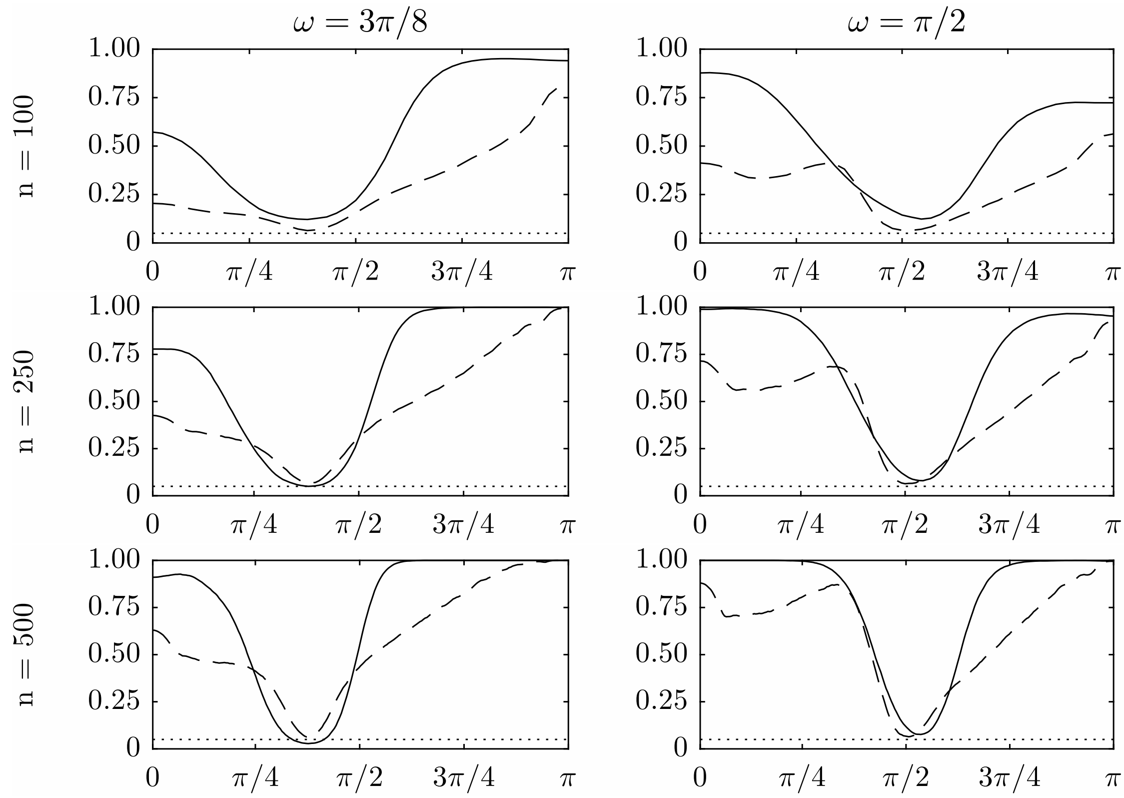
and stock prices for the US, studies long and short causality in a different aspect than ours. Furthermore, we explain how our method can be compared to them.

6.1 Term spread and real economic growth

A vast number of papers have showed the predictive power of interest rate spreads on real economic growth, among them Stock and Watson (1989), Bernanke (1990), Estrella and Hardouvelis (1991), Harvey (1991), Stock and Watson (1992), Estrella and Mishkin (1995), Bonser-Neal and Morley (1997), Davis and Fagan (1997), Hamilton and Kim (2002), Ang et al. (2006), Marcellino (2006), Galvão (2006).

More recently, Adrian et al. (2010) offered an alternative explanation of the Granger-causality of term spread on real economic growth. As pointed out by these authors, the traditional interpretation relies on the informational value of

Figure 2-5: Simulated power for DGP 4.ω with $\omega = \{3\pi/8, \pi/2\}$.



Notes: Solid line and dashed line represent the average rejection rate for our nonparametric test, $\hat{W}_\lambda$, and for B&C test, respectively. Dotted line represent the 5% rejection line. Sample size of {100, 250, 500} observations, $\nu = 1$, $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR order selected via AIC. 10,000 MC repetitions.

the yield curve for future short rates. However, they argue that the causal mechanism operates via financial intermediaries, including active management of balance sheets in response to changing economic conditions.

Following Breitung and Candelon (2006), we test the absence of causality between the term spread and the future economic growth. The term spread is defined by the difference between the government 10-year and the 3-month bond yield. We do not reject the null hypothesis, p-value = 0.032, at 1% of the augmented Dickey-Fuller test for the term spread. Thus we differentiated the series. The GDP growth rate is obtained with log-differences, $(100 \times \Delta \ln \text{real GDP}_t)$. In the second part of their empirical section, they follow Anderson and Vahid (2001) and investigate the impact of removing the influence of money growth, $(100 \times \Delta \ln \text{real M2}_t)$. The analyzed data comprehends the period from 1960.Q1 to 2017.Q3.

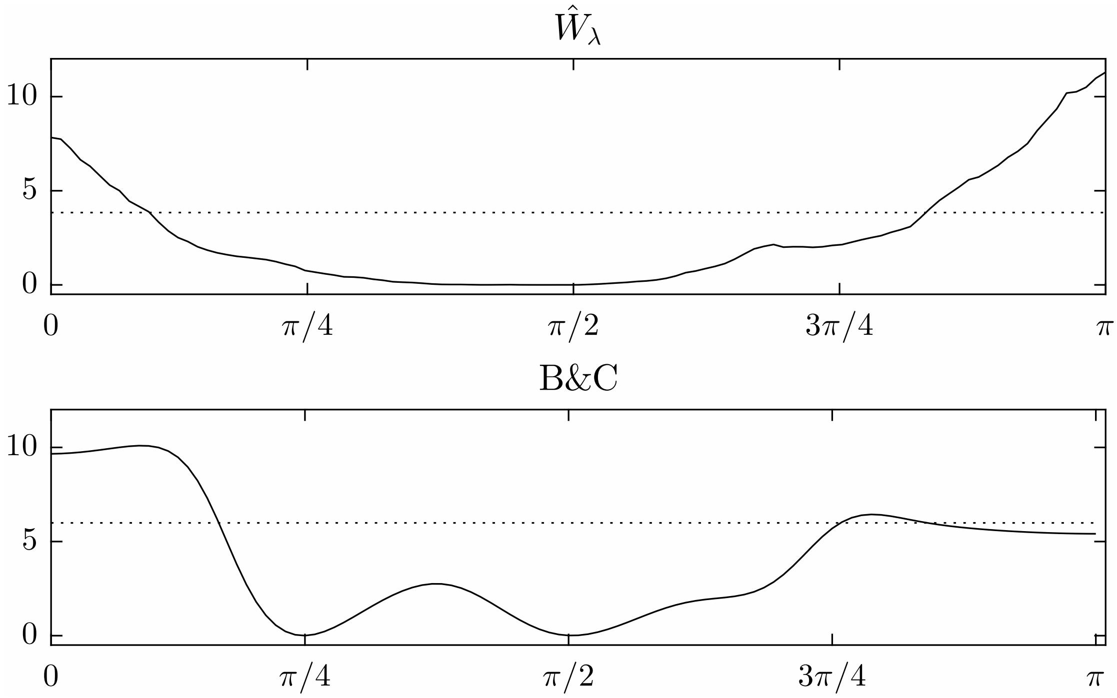
The results using our nonparametric causality test are similar to B&C test findings. Figure 2-6 reports the statistics of our nonparametric test and their test. In both cases, causality is reported at low frequencies, between $[0, 0.09\pi]$ for our nonparametric test and between $[0, 0.16\pi]$ for B&C test. Also, B&C report causality in the band spectrum $[0.76\pi, 0.83\pi]$ and our test on $[0.83\pi, \pi]$. The results did not change significantly with the inclusion of the money growth series, Figure 2-7. Following our test, the term spread still causing the real GDP growth for low frequencies, $[0, 0.07\pi]$, and additionally for the band spectrum $[0.83\pi, \pi]$. B&C reports causality in the $[0, 0.16\pi]$ and $[0.75\pi, 0.84\pi]$ band spectra. We also test the predictive power of the money growth on real GDP growth given the term spread, Figure 2-8. Neither our test nor B&C reports predictive power of M2 conditional to the term spread. According to Anderson and Vahid (2001), the term spread is a better leading indicator of output growth than M2. Our findings corroborate with their results.

6.2 Stock market and macroeconomic policy variables

Extending Hsiao (1982) and Lütkepohl (1993) work, Dufour and Renault (1998) and Dufour et al. (2006) proposes a generalization and a methodology to test the causality effect at an arbitrary horizon. Employing this technique Dufour and Tessier (2006) test the predictability of a set of macroeconomic policy variables and stock prices for the US. The used dataset comprehends GDP, Monetary Base, M1 multiplier, CPI, S&P 500 index, and Federal funds rate. All series, except for the Federal funds rate, are taken in logarithm. Following the guidelines of Dufour and Tessier (2006), we differentiate all series to transform in to stationary series. The series were not seasonality adjusted, according to Dufour and Tessier (2006) Granger-causality relies on projections of one series to another and the filtering process, e.g., X-11 filter, may produce unintended interference on results.

For GDP, Dufour and Tessier (2006) identify causality only from the Federal funds rate at different horizons. Here, we confirm their findings but also identify predictability power in the long and short run from S&P 500 and in the long run for CPI. For CPI, they found influence from GDP over all horizons. Our results suggest that not GDP but the interest rates are the primary predictor variable for prices. GDP seems to only affect prices in the long run. For interest rates, GDP is considered by Dufour and Tessier (2006) results as the primary source of predictive power. Our results also share this outcome. Furthermore, we also report the presence of short-run causality for Monetary Base, M1 multiplier, and S&P 500. Finally, for S&P 500 Dufour and Tessier (2006) did not detect any source of causality. However, we identify that for the long run no variable present predictive power but for the transition of long and short run and for short run, all series, except for GDP, causes the stock market index. This result is impressive since shows that markets are unaffected by macroeconomic policy variables in the

Figure 2-6: Test statistics for spectral representation of Granger-causality of term spread on real economic growth. Dataset comprehends the period from 1960.Q1 to 2017.Q3.



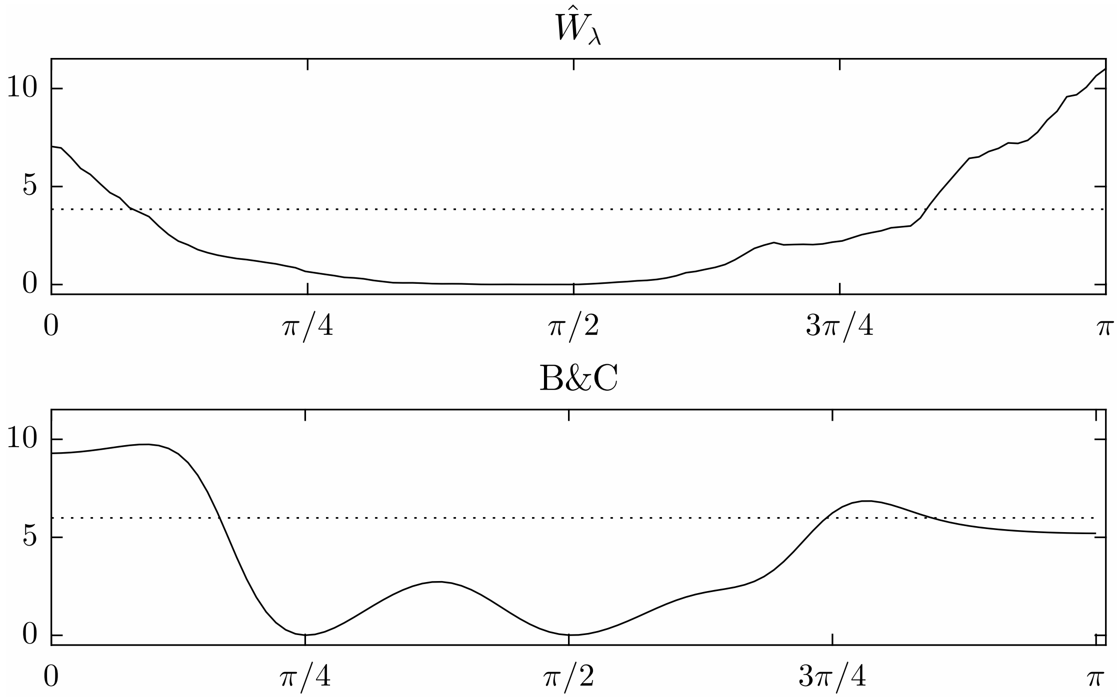
Notes: Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}] = 11$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR(9) selected via AIC. Dotted line represent the 5% rejection line.

long run but affected in the short run. The Efficient-market hypothesis, Fama (1970), is not contradicted by our results and sheds some light on markets short-run predictability theories, see Fama (1991).

Regarding multivariate causality, at least one of the variables cause GDP, CPI, and the interest rate at every frequency, which means that we have predictive power over the entire spectrum. As mention before, the set of macroeconomic policy variables does not predict S&P 500 in the long run. The non-predictability period goes from frequency 0 to 0.19π, which represent a time span from around two years and a half to infinity.

As in Dufour and Tessier (2006), the Monetary Base and the M1 multiplier did not present a clear interpretation. Despite some causality reports, as S&P 500 showing to be a predictor of Monetary Base in the short run, the Monetary Base and the M1 multiplier, in general, does not show, or show a marginal, linear dependence with the other variables.

Figure 2-7: Test statistics for spectral representation of Granger-causality of term spread on real economic growth given money stock (M2). Dataset comprehends the period from 1960.Q1 to 2017.Q3.



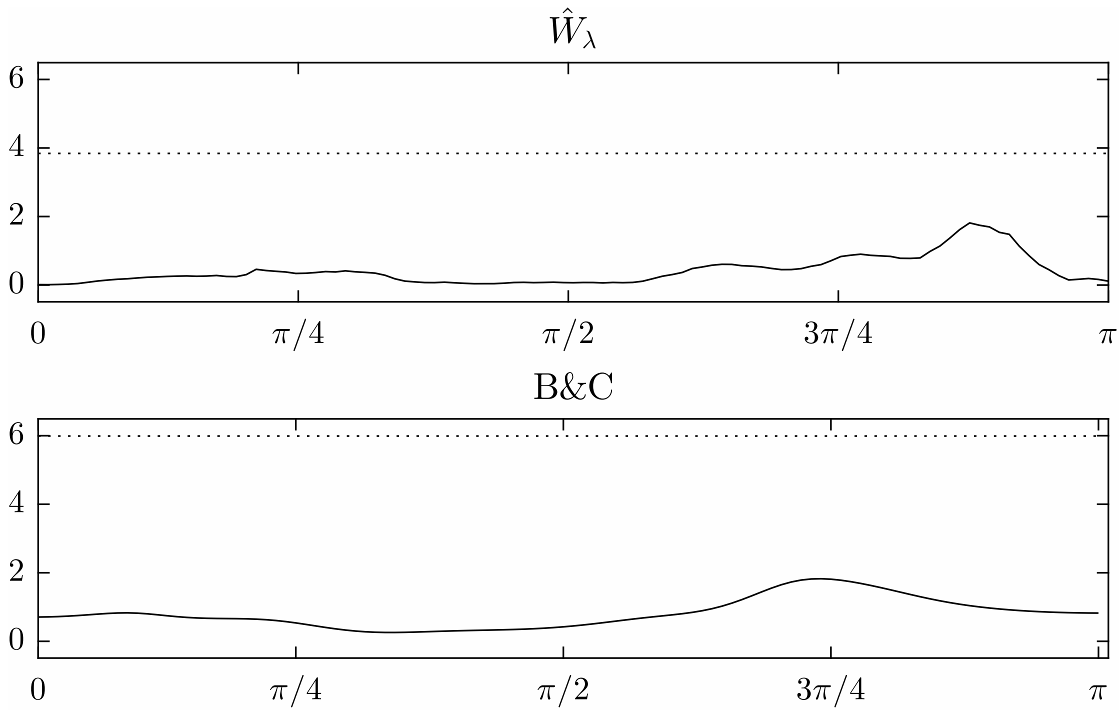
Notes: Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}] = 11$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR(9) selected via AIC. Dotted line represent the 5% rejection line.

7 Conclusion

We have proposed a novel nonparametric causality test, based on a two-step procedure, to infer about the non-existence of Granger-causality from a series to another at a given frequency. In the first step, we remove any inverse causality effect resulting from the presence of y_t lags. In the second step, using nonparametric estimation, we present a test that under the null converges to a χ^2 distribution. The test is easy to implement, and it is computationally inexpensive. The codes used in the experiments are available under request.

Monte Carlo experiments showed that our method has right size and power for a variety of models, and it is comparable with results using the parametric test of Breitung and Candelon (2006). Furthermore, our test showed a better response to causality under smooth coefficients over the spectrum. Empirical results using our method indicate similar conclusions to those using Breitung and Candelon (2006) method. B&C and our test suggest that low frequencies of the term spread

Figure 2-8: Test statistics for spectral representation of Granger-causality of money stock (M2)on real economic growth given term spread. Dataset comprehends the period from 1960.Q1 to 2017.Q3.

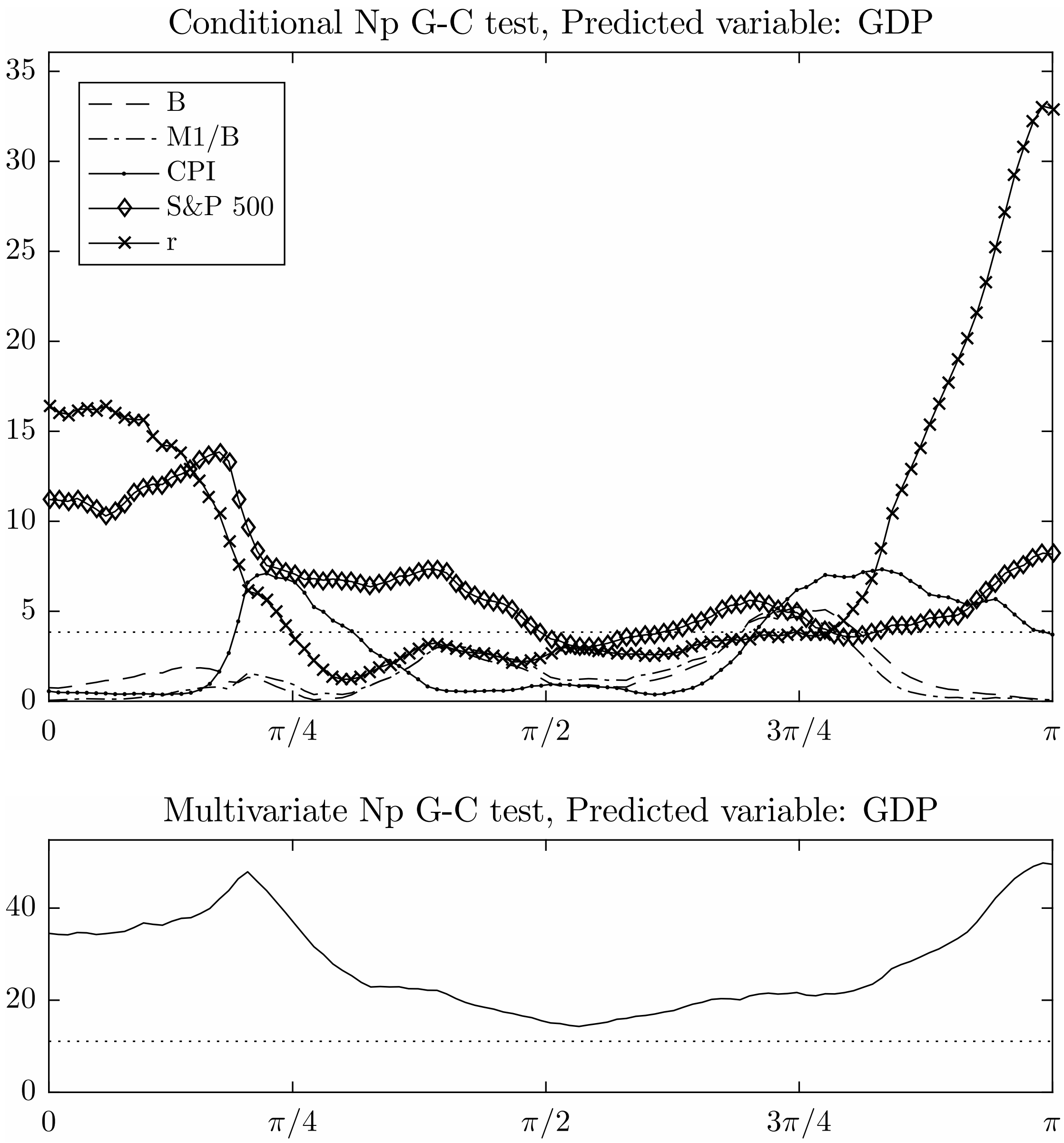


Notes: Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}] = 11$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. B&C test was performed up to a VAR(q), with VAR(9) selected via AIC. Dotted line represent the 5% rejection line.

to cause the GDP growth. The addition of a third variable, real money grown, does not alter the results significantly. None of the tests found causality from real money grown to real GDP grown given the term spread. We also reinterpreted the results of Dufour and Tessier (2006) of long and short run causality analyzing the behavior of their same dataset over the spectrum.

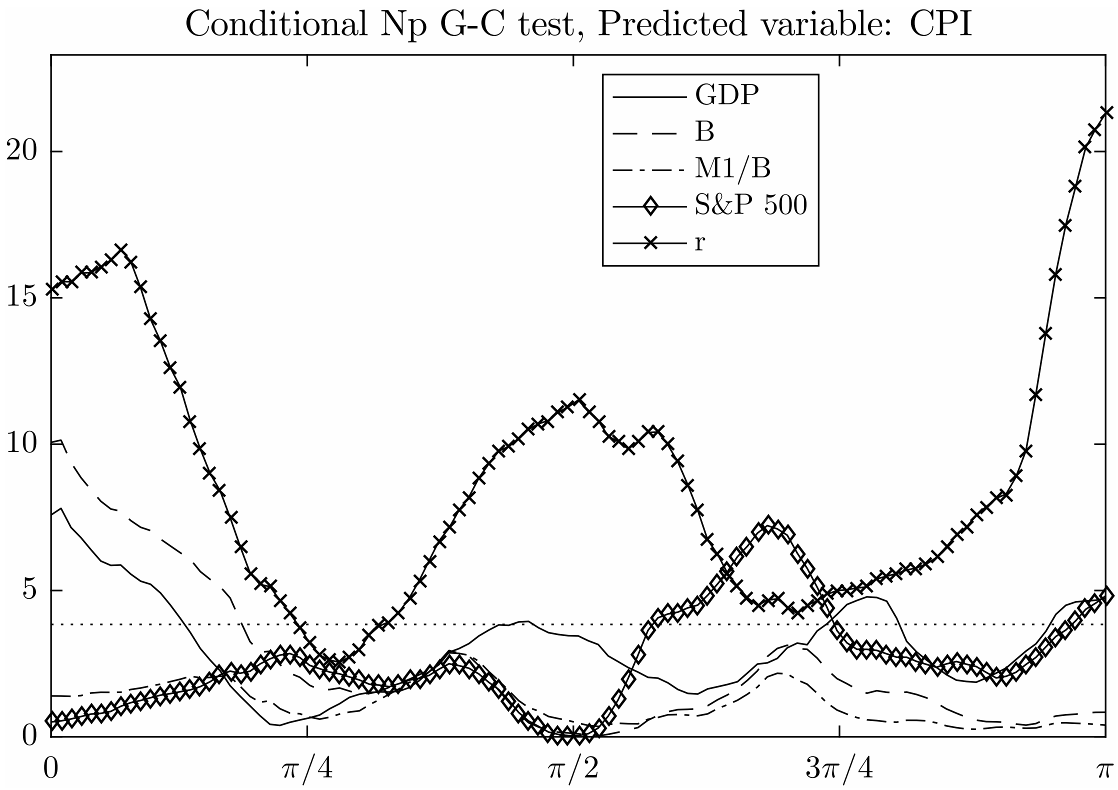
In addition to the statistical novelty, our method presents some interesting features as not to be conditioned to a specific kind of model, neither requires heteroskedasticity and autocorrelation consistent variance estimator for series with conditional variance. Furthermore, compared with Breitung and Candelon (2006) test, our nonparametric model reports an overall higher power and a good size performance. It would be worth to investigate, as future work, how our test can be accommodated in the test framework of Breitung and Schreiber (2017), i.e. testing if there at least one frequency in a band spectrum $\mathscr{B}$ such that $g_{\vec{xy}}(\lambda) = 0$ and the behavior of the phase shift when causality is present.

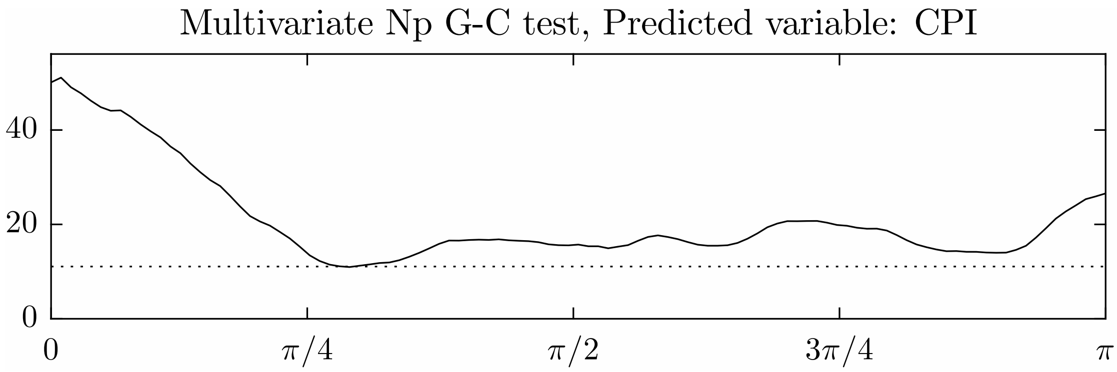
Figure 2-9: Test statistics for spectral representation of Granger-causality between Monetary Policy Variables and Stock Prices, with GDP as dependent variable. Dataset comprehends the period from 1959.Q1 to 2015.Q3.



Notes: Predictors: Monetary Base (B), M1 multiplier (M1/B), Consumer Price Index (CPI), Stock Price index (S&P 500), and Federal Funds rate (r). Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. Dotted line represent the 5% rejection line.

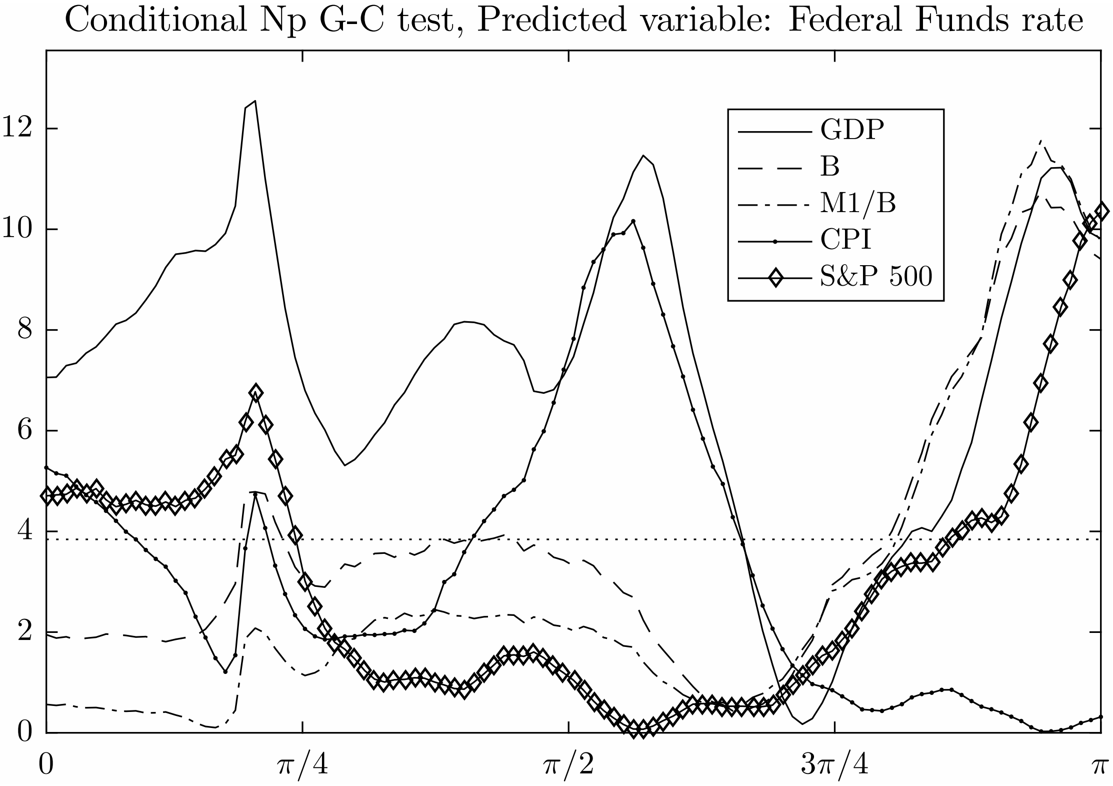
Figure 2-10: Test statistics for spectral representation of Granger-causality between Monetary Policy Variables and Stock Prices, with CPI as dependent variable. Dataset comprehends the period from 1959.Q1 to 2015.Q3.

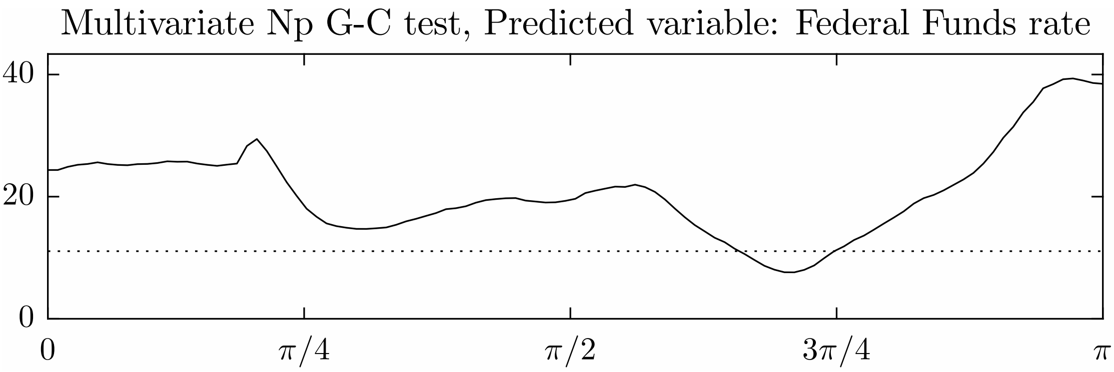




Notes: Predictors: Gross Domestic Product (GDP), Monetary Base (B), M1 multiplier (M1/B), Stock Price index (S&P 500), and Federal Funds rate (r). Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. Dotted line represent the 5% rejection line.

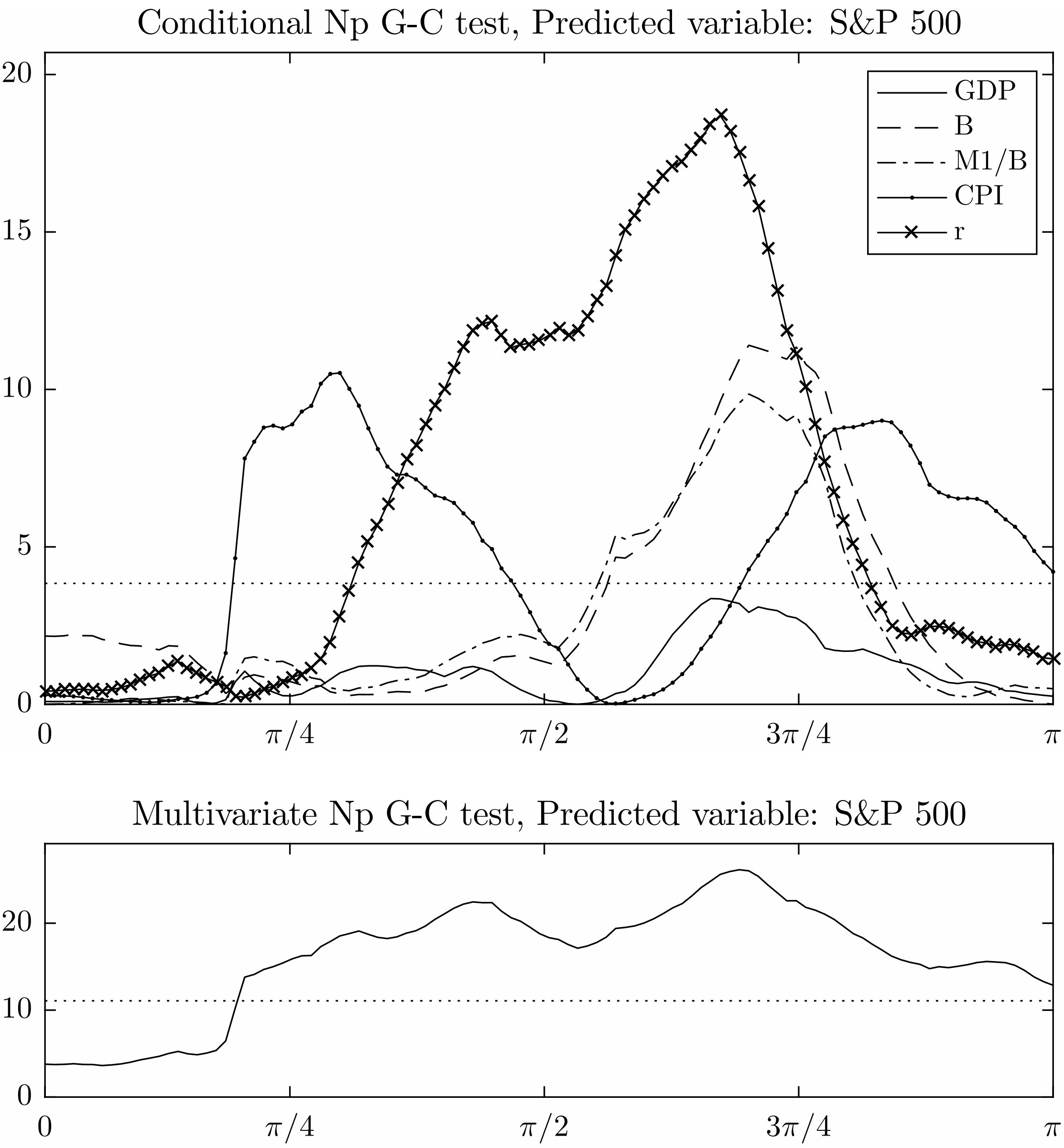
Figure 2-11: Test statistics for spectral representation of Granger-causality between Monetary Policy Variables and Stock Prices, with Federal Funds rate as dependent variable. Dataset comprehends the period from 1959.Q1 to 2015.Q3.





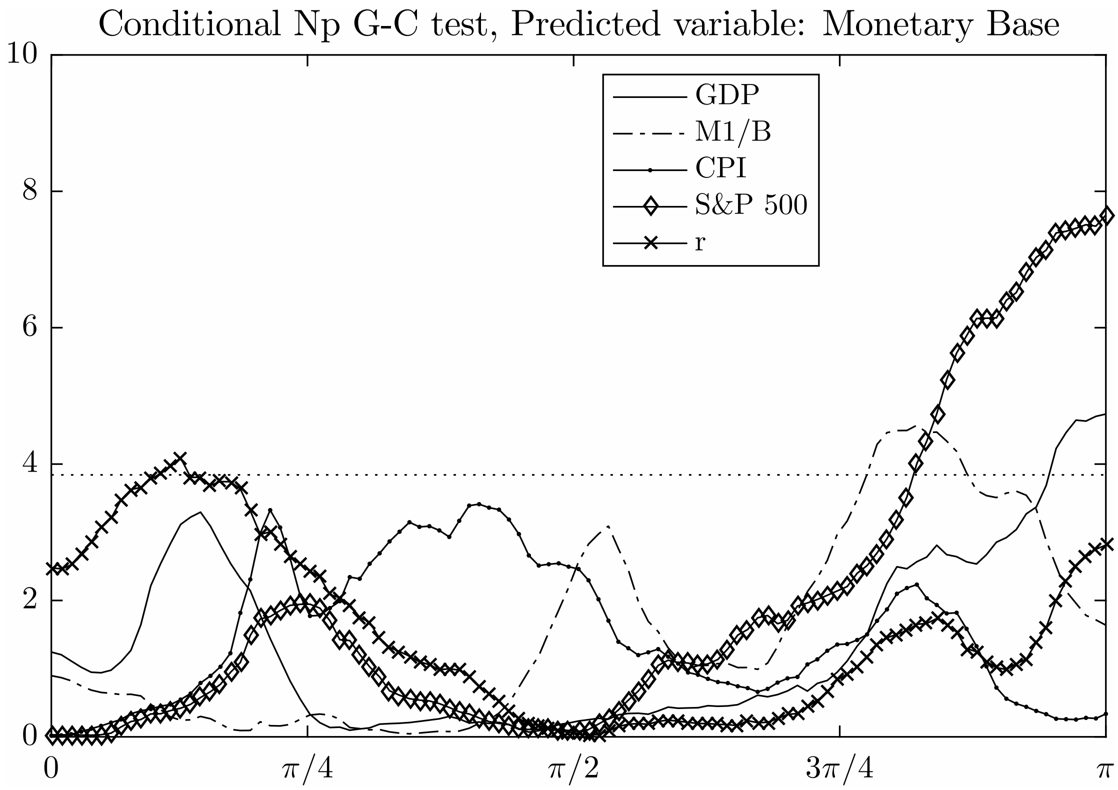
Notes: Predictors: Gross Domestic Product (GDP), Monetary Base (B), M1 multiplier (M1/B), Consumer Price Index (CPI), and Stock Price index (S&P 500). Nonparametric test, $\hat{W}_{\lambda}$, used $q = c_{\delta} n^{\delta} = [2n^{1/3}]$, $\tilde{n}h = c_{\gamma}\tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. Dotted line represent the 5% rejection line.

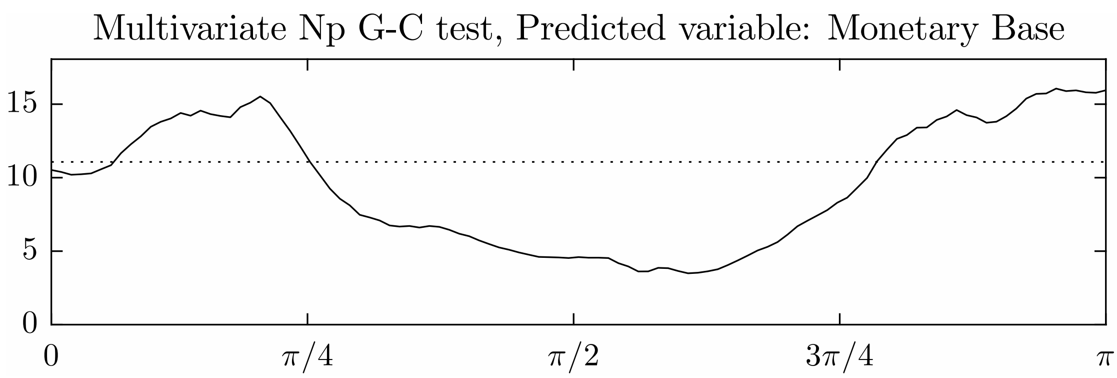
Figure 2-12: Test statistics for spectral representation of Granger-causality between Monetary Policy Variables and Stock Prices, with S&P 500 as dependent variable. Dataset comprehends the period from 1959.Q1 to 2015.Q3.



Notes: Predictors: Gross Domestic Product (GDP), Monetary Base (B), M1 multiplier (M1/B), Consumer Price Index (CPI), and Federal Funds rate (r). Nonparametric test, $\hat{W}_{\lambda}$, used $q = c_{\delta}n^{\delta} = [2n^{1/3}]$, $\tilde{n}h = c_{\gamma}\tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. Dotted line represent the 5% rejection line.

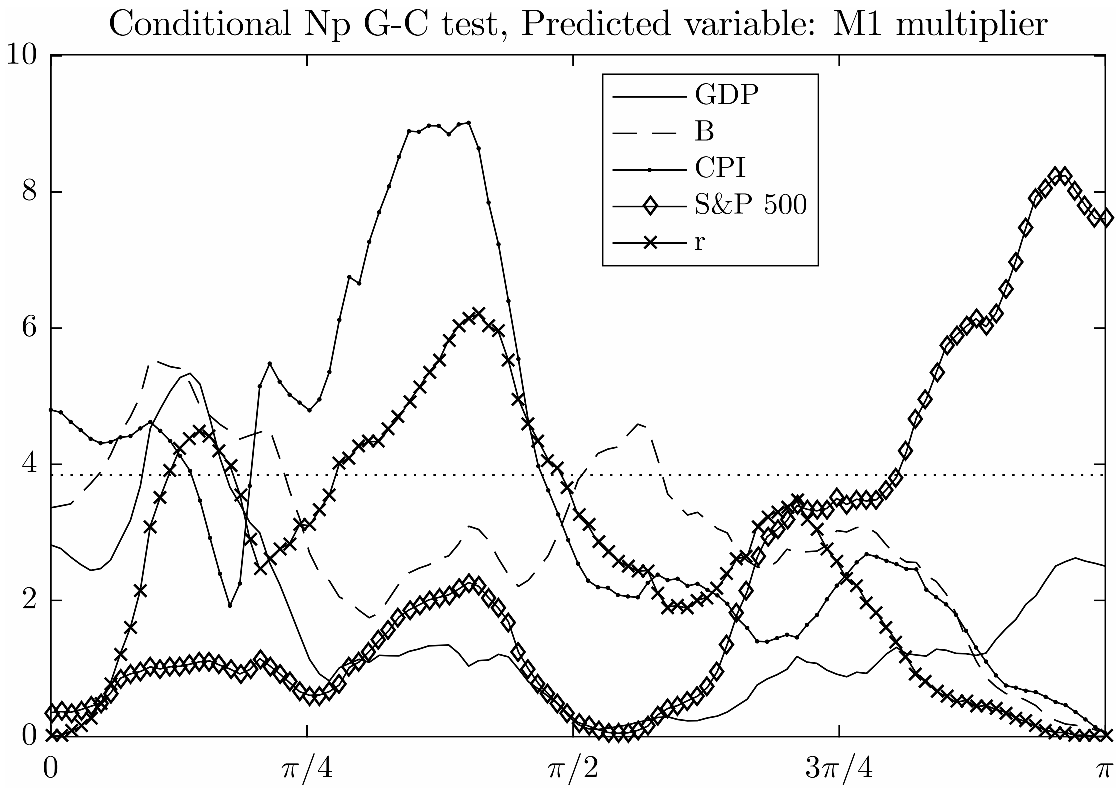
Figure 2-13: Test statistics for spectral representation of Granger-causality between Monetary Policy Variables and Stock Prices, with Monetary Base as dependent variable. Dataset comprehends the period from 1959.Q1 to 2015.Q3.

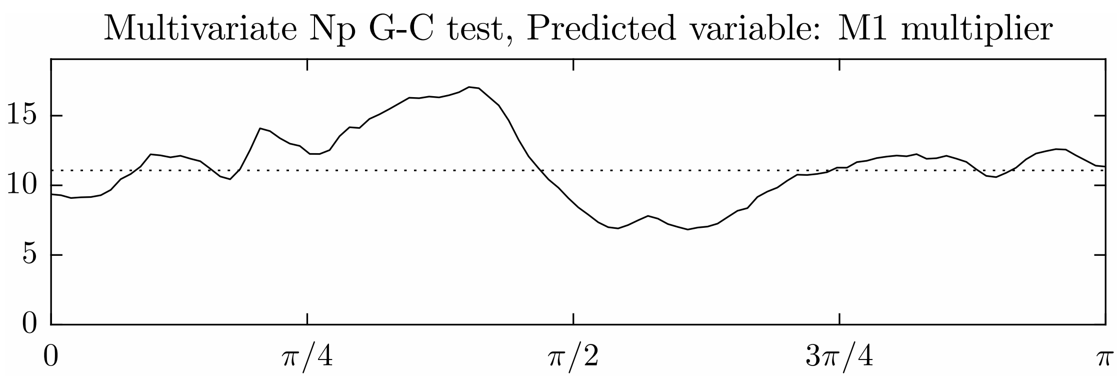




Notes: Predictors: Gross Domestic Product (GDP), M1 multiplier (M1/B), Consumer Price Index (CPI), and Stock Price index (S&P 500), and Federal Funds rate (r). Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. Dotted line represent the 5% rejection line.

Figure 2-14: Test statistics for spectral representation of Granger-causality between Monetary Policy Variables and Stock Prices, with M1 multiplier as dependent variable. Dataset comprehends the period from 1959.Q1 to 2015.Q3.





Notes: Predictors: Gross Domestic Product (GDP), Monetary Base (B), Consumer Price Index (CPI), Stock Price index (S&P 500), and Federal Funds rate (r). Nonparametric test, $\hat{W}_\lambda$, used $q = c_\delta n^\delta = [2n^{1/3}]$, $\tilde{n}h = c_\gamma \tilde{n}^{(1-\gamma)} \in [n^{0.45}, n^{0.55}]$ chosen by cross-validation. Dotted line represent the 5% rejection line.

Appendix 2.A

For a better understatement, we split the proof of Theorem 1 into three propositions. Also, to simplify notation on the proofs, let $\sum_{\lambda_s \in [-\pi,\pi]} = \sum_{\lambda_s}$.

Proposition 1. *Let $\hat{x}_t^0$ be obtained after the first step with x_t, $n \times p$, under Assumptions 3-7.*

$$\hat{f}_{\hat{x}_i^0 \hat{x}_j^0}(\lambda) = \frac{1}{\tilde{n}} \sum_{\lambda_s \in [-\pi,\pi]} K_{\tilde{n}h}(\lambda - \lambda_s) w^0_{\hat{x}_j}(\lambda_s) w^0_{\hat{x}_i}(\lambda_s)' \xrightarrow{p} f_{x_i^0 x_j^0}(\lambda).$$

Proposition 2. *Let y_t and x_t satisfy the Assumptions 3 to 7. Also, assume after the first step, that the estimation, defined in equation (2.5) on nonparametric model (2.4) respects Assumptions 6 and 7. Then, as $n \to \infty$ under $\mathbb{H}_0$,*

$$\frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s \in [-\pi,\pi]} K_{\tilde{n}h}(\lambda - \lambda_s) \left(\hat{\Theta}(\lambda) w_y(\lambda) + \Theta^+(\lambda) w_y(\lambda) \right) w^0_{\hat{x}}(\lambda_s)' \xrightarrow{p} 0.$$

and

$$\begin{aligned} \sqrt{h^{-1}} \hat{f}_{x^0 u^0}(\lambda) &= \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s \in [-\pi,\pi]} K_{\tilde{n}h}(\lambda - \lambda_s) w^0_{\hat{u}}(\lambda_s) w^0_{\hat{x}}(\lambda_s)' \\ &\xrightarrow{d} N\left(0, \eta(\lambda) f_{uu}(\lambda) f_{x^0 x^0}(\lambda) \int_{-\infty}^{\infty} K(s)^2 ds \right), \end{aligned}$$

with $\eta(\lambda) = 2$ for $\lambda = j\pi$, $j \in \mathbb{Z}$ and $\eta(\lambda) = 1$ otherwise. Also,

$$\hat{\Sigma}_{\overrightarrow{xy}}(\lambda) \xrightarrow{p} \eta(\lambda) f^{-1}_{x^0 x^0}(\lambda) f_{u^0 u^0}(\lambda) \int_{-\infty}^{\infty} K(s)^2 ds.$$

Proposition 3. *Given the results of Propositions 1-2, under $\mathbb{H}_0$ we have as $n \to \infty$,*

$$W_\lambda \xrightarrow{d} \chi^2_p, \quad W_{i,\lambda} \xrightarrow{d} \chi^2_1, \quad i = 1, \dots, p, \quad \lambda \in [0, \pi].$$

Proof of Proposition 1: According to Assumption 4, we have that $x_t = x_t^0 + \sum_{j=0}^{q} \phi_j y_{t-j}$, thus after first step we have:

$$\hat{x}_{t-1}^0 = x_{t-1}^0 + (\alpha_1 - \hat{\alpha}_1)y_{t-1} + \dots + (\alpha_q - \hat{\alpha}_q)y_{t-q} + \alpha_{q+1}y_{t-q-1} + \dots \tag{2.8}$$

with the number of observations dropping from n to $\tilde{n} = n - q = n - c_\delta n^\delta$. We can write equation (2.8) in frequency domain as,

$$\begin{aligned} w_{\hat{x}}^0(\lambda) &= w_x^0(\lambda) + (\phi_1 - \hat{\phi}_1)e^{-i\lambda}w_y(\lambda) + \dots + (\phi_q - \hat{\phi}_q)e^{-iq\lambda}w_y(\lambda) \\ &+ (\phi_{q+1}e^{-i\lambda} + \dots)w_y(\lambda) \\ &= w_x^0(\lambda) + \hat{A}(\lambda)w_y(\lambda) + A^+(\lambda_s)w_y(\lambda) \end{aligned}$$

for $\lambda \in [0, \pi]$ and with $\hat{A}(\lambda) = \sum_{j=1}^{q}(\alpha_j - \hat{\alpha}_j)e^{-ij\lambda}$, $A^+(\lambda) = \sum_{j=q+1}^{\infty} \alpha_j e^{-ij\lambda}$. In our framework, we have

$$\begin{aligned} \hat{f}_{x^0x^0}(\lambda) &= \frac{1}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) \left(w_x^0(\lambda_s) + \hat{A}(\lambda_s)w_y(\lambda_s) + A^+(\lambda_s)w_y(\lambda_s) \right) \\ &\times \left(w_x^0(\lambda_s) + \hat{A}(\lambda_s)w_y(\lambda_s) + A^+(\lambda_s)w_y(\lambda_s) \right)' \end{aligned}$$

then subdividing in three terms

$$\begin{aligned} \hat{f}_{x^0x^0}(\lambda) &= \frac{1}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) w_x^0(\lambda_s) w_x^0(\lambda_s)' \\ &+ \frac{2}{\tilde{n}} \operatorname{Re} \left[\sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) \left(\hat{A}(\lambda_s) + A^+(\lambda_s) \right) w_x^0(\lambda_s) w_y(\lambda_s)' \right] \\ &+ \frac{1}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) \left| \hat{A}(\lambda_s) + A^+(\lambda_s) \right|^2 w_y(\lambda_s) w_y(\lambda_s)' \end{aligned} \tag{2.9}$$

For the second term of RHS, we have that it is bounded by $2[\sup_{\lambda_s} |\hat{A}(\lambda_s)| + \sup_{\lambda_s} |A^+(\lambda_s)|](\tilde{n})^{-1} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) |w_x^0(\lambda_s) w_y(\lambda_s)'|$. By Theorem 2.1 and Lemma A.2 of Gonçalves and Kilian (2007), we know that $||\alpha_q - \hat{\alpha}_q|| = O_p(\frac{q^{1/2}}{n^{1/2}})$, and $\sup_{\lambda_s} |A^+(\lambda_s)| \to 0$ uniformly in λ by Assumption 4. We also have that $(\tilde{n})^{-1} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) |w_x^0(\lambda_s) w_y(\lambda_s)'|$ is $O_p\left((\tilde{n})^{-1} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda - \lambda_s) E||w_x^0(\lambda_s) w_y(\lambda_s)'|| \right) = O_p(1)$. In the third term of RHS, we have $\hat{f}_{yy}(\lambda) \xrightarrow{p} c > 0$, but the same arguments for $\hat{A}(\lambda_s)$ and $A^+(\lambda_s)$. Thus, second and third terms of (2.9) RHS converge to zero in probability.

Rosenblatt (1984) states that an extension of results to the multivariate case are straightforward. We, here, make some considerations of the extrapolation of these results, in special of Theorem 2 and Corollary 3 results. First, let us state a cross-spectrum estimate as in Rosenblatt (1984),

$$\bar{f}_{zr}(\lambda) = \frac{1}{2\pi} \sum_{j=-m}^{m} \frac{1}{n} \sum_{j=1}^{n} z_t r_{r+j} k_m(j) e^{-ij\lambda}.$$

with $1/m + m/n \to 0$, k_m satisfy Assumption 6 and $k_m(j) = k(j/m)$.

Robinson (1991b) shows, in Theorem 2.1, that both the lag-window, above, and the spectral-window, as 2.6, converge to the true spectrum value. Assumption 6 guaranties these results. Furthermore, in Stoica and Moses (2005) Chapter 2, an equivalence of the spectrum estimation through the cross-covariance function or through the smooth periodogram is established, see Robinson (1983) for the univariate equivalence. To see this, since $z(\lambda_s)'r(\lambda_s)$ is a cross-covariance estimator, with a suitable change in variables we have for (2.6),

$$\hat{f}_{zr}(\lambda) = \frac{1}{2\pi} \sum_{r=1-n}^{n-1} k\left(\frac{r}{nh}\right) e^{-i\lambda r} \hat{\Gamma}_{rz}(r) \tag{2.10}$$

with $\hat{\Gamma}_{rz} = n^{-1} \sum_{t=1}^{n-r} (r_t - \bar{r})(z_{t+r} - \bar{z})'$ and $\bar{z} = n^{-1} \sum_{t=1}^{n} z_t$. Whenever k has finite support, says $k(x) = 0$ for $x > 1$, the summation in (2.10) is restricted to $|r| < nh$, see Robinson (1991b).

Furthermore, our Assumptions 3-7 satisfy Corollary 3 of Rosenblatt (1984) and Theorem 2.1 of Robinson (1991b), thus $\hat{f}_{x^0x^0}(\lambda) \xrightarrow{p} f_{x^0x^0}(\lambda)$.

$\square$

Proof of Proposition 2: Notice that, $w_{\hat{y}}^0(\lambda) w_{\hat{x}}^0(\lambda)'$ under $\mathbb{H}_0$, can be written as

$$\left(w_u(\lambda) + \hat{\Theta}(\lambda) w_y(\lambda) + \Theta^+(\lambda) w_y(\lambda)\right) \left(w_x^0(\lambda) + \hat{A}(\lambda) w_y(\lambda) + A^+(\lambda) w_y(\lambda)\right)'$$

with $\hat{\Theta}(\lambda) = \sum_{j=1}^{q} (\theta_j - \hat{\theta}_j) e^{-ij\lambda}$, and $\Theta^+(\lambda) = \sum_{j=q+1}^{\infty} \theta_j e^{-ij\lambda}$.

Then, for the first part of Proposition 2, we have,

$$
\begin{aligned}
\frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} & K_{\tilde{n}h}(\lambda-\lambda_s)\left(\hat{\Theta}(\lambda_s)w_y(\lambda_s)+\Theta^{+}(\lambda_s)w_y(\lambda_s)\right)w^0_{\hat{x}}(\lambda_s)' \\
&= \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)\hat{\Theta}(\lambda_s)w_y(\lambda_s)w^0_{x}(\lambda_s)' \\
&+ \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)\hat{\Theta}(\lambda_s)w_y(\lambda_s)\left(\hat{A}(\lambda_s)w_y(\lambda_s)\right)' \\
&+ \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)\hat{\Theta}(\lambda_s)w_y(\lambda_s)\left(A^{+}(\lambda_s)w_y(\lambda_s)\right)' \qquad (2.11)\\
&+ \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)\Theta^{+}(\lambda_s)w_y(\lambda_s)w^0_{x}(\lambda_s)' \\
&+ \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)\Theta^{+}(\lambda_s)w_y(\lambda_s)\left(\hat{A}(\lambda_s)w_y(\lambda_s)\right)' \\
&+ \frac{\sqrt{h^{-1}}}{\tilde{n}} \sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)\Theta^{+}(\lambda_s)w_y(\lambda_s)\left(A^{+}(\lambda_s)w_y(\lambda_s)\right)'.
\end{aligned}
$$

We need to show that each element of 2.11 RHS goes to 0 in probability. The first term of RHS is majorized by $\sqrt{h^{-1}}\sup_{\lambda_s}|\hat{\Theta}(\lambda_s)|(\tilde{n})^{-1}\sum_{\lambda_s}K_{\tilde{n}h}(\lambda-\lambda_s)||w_y(\lambda_s)$ $w^0_x(\lambda_s)'||$. Notice that $(\tilde{n})^{-1}\sum_{\lambda_s}K_{\tilde{n}h}(\lambda-\lambda_s)||w_y(\lambda_s)w^0_x(\lambda_s)'||$ is $O_p(1)$ and by Theorem 2.1 and Lemma A.2 of Gonçalves and Kilian (2007) and n large enough we have that the second term of (2.11) RHS is bounded by

$$
\begin{aligned}
h^{-1/2}||\Theta_q-\hat{\Theta}_q||O_p(1) &= O_p\left(h^{-1/2}\frac{q^{1/2}}{n^{1/2}}\right) \\
&= O_p\left(n^{-\gamma/2}n^{(\delta-1)/2}\right) \\
&= o_p(1) \text{ if } \gamma+\delta<1.
\end{aligned}
$$

For the second term, we have that it is majorized by $\sqrt{h^{-1}}\sup_{\lambda_s}|\hat{\Theta}(\lambda_s)|sup_{\lambda_s}$ $|\hat{A}(\lambda_s)|\hat{f}_{yy}(\lambda)$. This implies in $\gamma+2\delta<2$ and as $\delta<1/3$, this condition becomes less restrictive than $\gamma+\delta<1$. By assumption, terms with + superscript have a faster rate of convergence than $\hat{A}$ or $\hat{\Theta}$, thus results follow. Note that bias in the kernel estimation is at order $(\tilde{n}h)^{-2}$, see Cai (2007), thus,

$$
h^{-1/2}(nh)^{-2}=n^{\gamma/2}n^{-2+2\gamma}=n^{(5\gamma-4)/2}\rightarrow 0 \text{ because } \gamma<4/5 \text{ by Ass. 5.}
$$

For the second part, we have,

$$\sqrt{h^{-1}}\hat{f}_{\hat{x}^0 u}(\lambda) = \frac{\sqrt{h^{-1}}}{\tilde{n}}\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_u^0(\lambda)w_{\hat{x}}^0(\lambda)'$$

$$\begin{aligned} &= \frac{\sqrt{h^{-1}}}{\tilde{n}}\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_u(\lambda_s)w_x^0(\lambda_s)' \\ &+ \frac{\sqrt{h^{-1}}}{\tilde{n}}\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_u(\lambda_s)\hat{A}(\lambda_s)'w_y(\lambda_s)' \\ &+ \frac{\sqrt{h^{-1}}}{\tilde{n}}\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_u(\lambda_s)A^{+}(\lambda_s)'w_y(\lambda_s)', \end{aligned} \tag{2.12}$$

where the second term of (2.12) RHS is majorized by $h^{-1/2}||\alpha_q - \hat{\alpha}_q||\ \tilde{n}^{-1}\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)|w_{\hat{u}}(\lambda_s)w_y(\lambda_s)'| = h^{-1/2}||\alpha_q - \hat{\alpha}_q||O_p(1)$. Then, gathering previous results, the second and third term of (2.12) RHS $\xrightarrow{p} 0$. The first term of (2.12) can be written as

$$\frac{\sqrt{h^{-1}}}{\tilde{n}}\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_u^0(\lambda_s)w_x^0(\lambda_s)' + o_p(1), \tag{2.13}$$

then by Assumption 3-7, see Gonçalves and Kilian (2007) and Bühlmann (1995), conditions of Corollary 3 of Rosenblatt (1984) are satisfied. Thus, by pointwise estimation of the spectrum, (2.13) satisfies a CLT with variance described in Proposition 2 statement.

With respect to the third part of Proposition 2, gathering Proposition 1 and previous results of Proposition 2, we have that by the independence between $w_x(\lambda)$ and $u(\lambda)$, see Corbae et al. (2002), and for the convergence towards a unique frequency our desired result. □

Proof of Proposition 3: Gathering the results of Proposition 1-2, the desired result follows. □

Proof of Theorem 7: W.l.o.g. assume that x is univariate, by equation (2.4) we have that $g_{\vec{xy}}(\lambda) = f^{-1}_{\hat{x}^0\hat{x}^0}(\lambda)f_{\hat{x}^0\hat{y}^0}(\lambda)$, then

$$\sqrt{h^{-1}}\left(\hat{g}_{\vec{xy}}(\lambda) - g_{\vec{xy}}(\lambda)\right) = \sqrt{h^{-1}}\left(\frac{\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_{\hat{y}}^1(\lambda_s)w_{\hat{x}}^0(\lambda_s)'}{\sum_{\lambda_s} K_{\tilde{n}h}(\lambda-\lambda_s)w_{\hat{x}}^0(\lambda_s)w_{\hat{x}}^0(\lambda_s)'} - g_{\vec{xy}}(\lambda)\right)$$

we have that $w^1_{\hat{y}}(\lambda_s)$ is $(w^0_x(\lambda_s) + w_u(\lambda_s))$ plus a bias term. Thus, the term in respect $w^0_x(\lambda_s)$ is majored by

$$\sqrt{h^{-1}}\left(g_{\vec{xy}}(\lambda_s) - g_{\vec{xy}}(\lambda)\right)(1 + o_p(1)) \leq \sqrt{h^{-1}}\sup_{\lambda_s}||g_{\vec{xy}}(\lambda_s) - g_{\vec{xy}}(\lambda)||\,(1 + o_p(1))$$

which is $O_p(h^{1/2})$, since $\max|\lambda - \lambda_s| = 2\pi(j \pm \tilde{n}h)/\tilde{n} = 2\pi j/\tilde{n} \pm 2\pi h \to 0$.

The term involving $w_u(\lambda_s)$ results in

$$(f_{x^0x^0}(\lambda) + o_p(1))^{-1}\sqrt{h^{-1}}f_{x^0u}(\lambda) + o_p(1)$$

which according Proposition 2 and 3 converges to $\Sigma_{\vec{xy}}(\lambda)$.

The covariance matrix Σ is given by Theorem 6 proof. For the second part of Theorem 7, we have

$$W_\lambda = \left(g_{\vec{xy}}(\lambda) + o_p(1)\right)^2 \frac{(f_{x^0x^0}(\lambda) + o_p(1))}{(f_{uu}(\lambda) + o_p(1))}h^{-1} = g^2_{\vec{xy}}(\lambda)\frac{f_{x^0x^0}(\lambda)}{f_{uu}(\lambda)}h^{-1}\,(1 + o_p(1))$$

then, since $f_{x^0x^0}(\lambda)$ and $f_{uu}(\lambda) > 0$, take $c = g^2_{\vec{xy}}(\lambda)f_{x^0x^0}(\lambda)f^{-1}_{uu}(\lambda)$. The result for $W_{i,\lambda}$ follow the same lines and it is omitted.

$\square$

Chapter 3

Nonparametric causality test for mixed-frequency datasets

Abstract

We propose two novel nonparametric causality tests for mixed-frequency datasets. One based on least squares, LS, estimation and other based on the Hannan-Inefficient, HI, estimator. In our framework, the dependent variable is fitted by an increasing, with the sample size, number of leads and lags of the exogenous variable. The LS approach presents better results for testing causality at individual leads/lags, and the HI approach is superior for addressing the joint null hypothesis of non-causality from one series to another. Furthermore, we assume that the low-frequency variable is generated at high frequency, but some observations are systematically unobserved, rather than assume distinct generation frequencies as in MIDAS and mixed-frequency VAR models. Thus, our approach results in a more straightforward interpretation of causality. Finite sample simulations show good results in size and power for our causality tests and consistent estimation of leads and lags coefficients. Finally, we provide some empirical results on testing for no causality between GDP and US monthly indicators.

Keywords: Causality, Frequency-domain, Nonparametric test, Mixed-Frequencies.

JEL classification: C12, C14, C32.

1 Introduction

Mixed-frequency datasets contain series that were sampled at different frequencies. For example, let $\{y_t, x_t\}$ be two stationary series with n observations, $n \mod s$, and s be the subsampling ratio. Series $y_{s,t}$ is defined as the subsampled version of y_t, with one observation ever s observations of x_t. For simplicity, henceforth, we assume that only two sample rates are present in a dataset.

Here, we propose two nonparametric causality testing approaches based on a two-sided unrestricted distributed-lag model, i.e., a distributed-lag model with an infinite number of leads and lags. The first method is based on least squares, LS, and the second is based on the Hannan-Inefficient, HI, estimator, see Hannan (1963b). In both cases, using lag coefficients, we can infer the absence of causality from the past values of x_t to $y_{s,t}$, and using lead coefficients, from past values of $y_{s,t}$ to x_t, as in Sims (1972). We also investigate causality at a specific lead or lag. For coefficient estimation purpose we assume that number of leads and lags, M, increases with n but slowly.

The so-called 'Hannan-Inefficient', HI, estimator, Hannan (1963b), receives its name from Sims (1973) in comparison to the fully efficient estimation also proposed in Hannan (1963b). Sims argues that for the univariate case this procedure is equivalent asymptotically to GLS. According to Sims (1973), the attractiveness of HI estimator over the LS regression shows up whenever the estimation of a large number of lags is necessary. Furthermore, seasonal adjustments are easy to handle, and it automatically handles serial correlation in the residuals. See Amemiya and Fuller (1967), Hannan (1967), and Wahba (1969) for further studies on the HI estimator.

Wahba (1969) studied the asymptotic properties of HI estimator for a two-sided unrestricted distributed-lag model. Based on a similar framework, Hidalgo (2000, 2005) proposed a nonparametric test for Granger-causality that is consistent for stationary series presenting long memory. His test assures to measure the causality absence for the whole series, say x_{t-j} does not cause y_t for $\forall j \geq 1$.

Extending Hsiao (1982) and Lütkepohl (1993) works, Dufour and Renault (1998) and Dufour et al. (2006) propose a generalization and a methodology to test the causality effect at an arbitrary horizon h. One can interpret measurements of causality at different horizons, see Dufour and Taamouti (2010), as an alternative to frequency domain approaches such Breitung and Candelon (2006) or Taufemback (2018b). For example, a non-zero effect concerning a lag approaching infinity represents a long-run causality in the time domain and low-frequency causality in the frequency domain. Here, we develop our methodology on the time domain but also using spectral estimates.

Recently, Ghysels et al. (2016) and Götz et al. (2016) propose two similar Granger-causality test for mixed-frequency datasets. The first one is more indicated for a small difference in the subsampling rate and the second for high values of s. Based on mixed-frequency VAR models, both tests suffer from parameter proliferation as s increases. The first uses parametric bootstrap and the second Bayesian estimation as techniques for parameter reduction. These tests also allow for testing causality from the low-frequency, LF, variable to the high-frequency, HF, one.

Following Taufemback (2018a), we assume that the series, $y_{s,t}$ is generated at the same frequency of x_t, but some observations are systematically unobserved. Mariano and Murasawa (2003, 2010), and Schorfheide and Song (2015) also employed this data generation structure. However, in contrary, MIDAS and mixed-frequency VAR methods rely on mixed frequency data generation, see Ghysels et al. (2004b, 2007b); Ghysels (2016). For example, in a quarterly-monthly dataset, the first variable is generated at a frequency three times slower than the last one. For more details about these methods see Guérin and Marcellino (2013); Foroni and Marcellino (2013); Götz et al. (2014); Foroni et al. (2015b); Marcellino and Schumacher (2010); Bai et al. (2013b); Foroni et al. (2015a); McCracken et al. (2015).

Some notations used through the paper: s means the ratio between the number of observations of high frequency, n, and low frequency variables, n_s; $z_{s,t}$, $n_s \times p$, represents the low frequency version of a high frequency variable z_t, $n \times p$. The variable Z_n represents the vector $\{z_t\}_{t=1}^{n}$, and $Z_{s,n}$ represents the vector $\{z_t\}$ sampled at $t = s, 2s, \dots, n$. $w_z = WZ_n$, where W, $n \times n$, represents the discrete Fourier transform with row j given by $W_j = n^{-1/2}(1, e^{-i\lambda_j}, e^{-i2\lambda_j}, \dots, e^{-i(n-1)\lambda_j})$, $\lambda_j = 2\pi j/n$, $j = 0, \dots, n-1$. Also, let $w_{z_s} = W_s Z_{s,n}$, where W_s is defined as W, but for low frequency variables. We represent that z_{t-j} does not cause r_t for a specific j^* as $z_{t-j^*} \not\to r_t$, i.e., $E[r_t|r_{t-l}, z_{t-j}, l > 0, j > 0] = E[r_t|r_{t-l}, z_{t-j}, l > 0, j > 0, j \neq j^*]$, and $Z_n \not\to R_n$ means that any past values of z_t cause r_t, $E[r_t|r_{t-j}, z_{t-j}, j > 0] = E[r_t|r_{t-j}, j > 0]$. The prime symbol, as in $w_z(\lambda)'$, means transpose conjugated.

The organization of the article is the following. Section 2 discusses mixed-frequency causality and our least squares proposal. Section 3 present a mixed-frequency Hanan-Inefficient estimator. Section 4 studies the assumptions and the asymptotic theory. Section 5 presents some considerations about finite sample size estimation. Section 6 present some Monte Carlo experiments designs. Accommodating four mixed-frequency dataset formats: quarterly-monthly, annually-quarterly, annually-monthly, and monthly-daily. Section 7 presents empirical results. Section 8 concludes the paper. Appendix A presents the functional form of some matrices, Appendix B theorems proofs, and Appendix C reports Monte Carlo simulation results.

2 Mixed-Frequency Causality test

According to Chamberlain (1982), see also Kuersteiner (2010), the Granger causality that X_n does not cause Y_n can be described as

$$\text{(G) } y_{t+1} \text{ is independent of } x_t, x_{t-1}, \ldots, \text{ conditional on } y_t, y_{t-1}, \ldots, \forall t$$

however, Sims (1972) shows that if causality only occurs in one direction, say from the past values of x_t to y_t, then an augmented regression with x_t leads will result in null coefficients for those extra variables. Also according to Chamberlain (1982), the Sims's causality definition, or the strict exogeneity condition, is given by

$$\text{(S) } x_t \text{ is independent of } y_{t+1}, y_{t+1}, \ldots, \text{ conditional on } y_t, y_{t-1}, \ldots, \forall t$$

We explore these definition to measure causality between y_t and x_t, in both directions, in a two-sided unrestricted distributed-lag model. So, let y_t and x_t be two zero mean strict stationary time series with a linear relationship in the time domain, as in Wahba (1969) and Hidalgo (2000), described by,

$$y_t = \sum_{j=-\infty}^{\infty} c(j)x_{t-j} + u_t, \; E[u_t|x_s, -\infty < s < \infty] = 0, \; \sum_{j=-\infty}^{\infty} ||c(j)|| < \infty \quad (3.1)$$

where $c(j)$ is a $1 \times p$ vector. Thus, we define our null hypotheses of joint non-causality by,

$$\begin{aligned} &\mathbb{H}_0(X_n \nrightarrow Y_n) : \forall j > 0, \; c(j) = 0 \;\text{ with }\; \mathbb{H}_1(X_n \rightarrow Y_n) : \exists j > 0 \text{ s.t. } c(j) \neq 0, \\ &\mathbb{H}_0(Y_n \nrightarrow X_n) : \forall j < 0, \; c(j) = 0 \;\text{ with }\; \mathbb{H}_1(Y_n \rightarrow X_n) : \exists j < 0 \text{ s.t. } c(j) \neq 0, \end{aligned}$$

in other words if the sum of absolute value of lagged x_t coefficients is equal to zero then past values of x_t does not cause y_t. Conversely, if the sum of the absolute coefficients concerning to the leads of x_t is zero then we say that past values of y_t do not cause x_t.

Four cases of causality may occur, $\{\{X_n \nrightarrow Y_n, Y_n \nrightarrow X_n\}, \{X_n \rightarrow Y_n, Y_n \nrightarrow X_n\}, \{X_n \nrightarrow Y_n, Y_n \rightarrow X_n\}, \{X_n \rightarrow Y_n, Y_n \rightarrow X_n\}\}$. Comparing our representation (3.1) with traditional models shows a drawback only for the third case. For example assume that y_t and x_t are generated as $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + v_{1,t}$ and $x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \cdots + v_{2,t}$, for independent iid series v_{1t} and v_{2t}, which implies that $Y_n \rightarrow X_n$ but $X_n \nrightarrow Y_n$. Now, representing these DGPs as in (3.1) will report, apart from causality from $Y_n \rightarrow X_n$, spurious causality from X_n to Y_n whenever we have $\alpha_j \neq 0$, for $j \geq 1$, since is possible to show, for this case, that $c(j)$ will be different from zero for $j > 0$. Thus, we need to be careful in how we report our results for this case. Unfortunately, a similar approach to Taufemback (2018b), i.e., perform a prewhitening on y_t in respect to its lags, cannot be employed for mixed-frequency datasets since we do not observe all realizations of y_t.

Now, for a strict stationary mixed frequency dataset, let the observations of x_t be obtained at a sampling frequency normalized to 1 observation at some unit of time, and observations of y_t be only obtained at sampling frequency $\omega_s = 1/2s$, for some integer s. The subsampled series of y_t is designed as $y_{s,t}$ with $n_s = \lfloor n/s \rfloor$ observations. Henceforth, we assume that n is mod s, which implies that x_t has sn_s, observations.

Despite we are working with mixed-frequency dataset, least squares for the estimation of $c(j)$ remains trivial. Let M be the number of fitted leads and lags and $\mathcal{X}_t^M = (x_{t+M}, \ldots, x_t, \ldots, x_{t-M})$, and define $\mathcal{X}_{s,n}^M$ as the temporal aggregated version of the matrix $\{\mathcal{X}_t^M\}_{t=1}^n$ with dimensions of $n \times (2M+1)p$. Thus, the least squares estimation of $c_M = \{c(-M), \ldots, c(M)\}$ is given by

$$\hat{c}_M = (\mathcal{X}_{s,n}^{M\prime} \mathcal{X}_{s,n}^M)^{-1} \mathcal{X}_{s,n}^M Y_{s,n}, \tag{3.2}$$

where we assume that M grows with n, but at lower rate. With the exception of the leads, this model is similar to U-MIDAS, Foroni et al. (2015b) with finite number of lags. The null hypotheses of joint causality for mixed-frequency is the same from regular datasets.

Notice that one could test with all leads/lags of an specific variable in X_n, say X_n^p, do not cause Y_n. The null hypothesis follows the same framework described above and we further refer to this test as $\mathbb{H}_0(X_n^p \not\to Y_n)$. We also define as $\mathbb{H}_0^j(X_n \not\to Y_n)$ as the null hypothesis in terms an specific lead/lag j.

Causality tests with mixed-frequency VAR, see Ghysels et al. (2016) and Götz et al. (2016), demand the estimation of a large matrix of coefficients. This approach requires stack the low-frequency and high-frequency variables in a vector, say $Z_t = (x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-s}, y_{s,t})'$, then estimate a reduced VAR, e.g. $Z_t = AZ_{t-1} + u_t$. As example, for a mixed-frequency VAR(1) we have,

$$\begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-s} \\ y_{s,t} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,s+1} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,s+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s,1} & a_{s,2} & \cdots & a_{s,s+1} \\ a_{s+1,1} & a_{s+1,2} & \cdots & a_{s+1,s+1} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \\ \vdots \\ x_{t-s-1} \\ y_{s,t-s} \end{bmatrix} + u_t, \tag{3.3}$$

thus, $x_{t-j} \not\to y_{s,t}$ if $a_{s+1,j} = 0$, for $j = 1, \ldots, s$, and the inverse causality does not occurs if and only if $a_{j,s+1} = 0$, $\forall j = 1, \ldots, s$.

Mixed-frequency VAR models assume that the series are generated at different frequencies. In a quarterly-monthly dataset, the first past value of the quarterly series is given by the last quarter, say $t-s$. The value of y_{t-1} is inexistent. Under our framework, we assume that these variables exist but are unobservable. Furthermore, our set up implies that if $c(-1) \neq 0$ then $y_{s,t-1} \to x_t$.

3 Mixed-Frequency Hannan-Inefficient Estimator

An alternative to the least squares estimator of $c(j)$ is the Hannan-Inefficient, HI, estimator. Suggested by Hannan (1963b), the HI estimator has interesting characteristics due its simplicity and flexibility. Hidalgo (2000) proposes explore the characteristics of HI estimator to overcome least squares issues with long memory, see Robinson (1994). The HI estimator for regular datasets is given by,

$$\check{c}(j) = \frac{1}{2M} \sum_{q=0}^{2M-1} \hat{f}_{xx}^{-1}(\lambda_{2mq}) \hat{f}_{xy}(\lambda_{2mq}) e^{ij\lambda_{2mq}}, \quad \lambda_l = 2\pi l/n, \tag{3.4}$$

where $j \in \mathbb{Z}$ and

$$\hat{f}_{zg}(\lambda) = \frac{1}{2m+1} \sum_{k=-m}^{m} w_g(\lambda + \lambda_k) w_z(\lambda + \lambda_k)', \quad m = [n/4M], \tag{3.5}$$

is the estimate of the cross-spectral density of z_t and g_t at frequency λ.

Hidalgo (2000) argues that for testing $\mathbb{H}_0(X_n \not\to Y_n)$, based on $\check{c}(j)$ estimates with a Wald test, M must increase slowly to n. Otherwise, the implementation of such test would be not straightforward. His suggestion for this case is the use of Lütkepohl and Poskitt (1996) upper bound, i.e., $M = O(n^{1/3})$.

In contrary to the simplicity of least squares, we need to discuss some results to implement a mixed-frequency HI, MF-HI, estimator. As in Taufemback (2018a), we assume that series y_t is only observable for $t = \{s, 2s, 3s, \dots, n\}$. Which implies not only a re-sample procedure but also a shift in time. Lemma 7 defines the relationship between y_t and $y_{s,t}$ in frequency domain.

Lemma 7 (Taufemback (2018a)). *Let (w_y, w_{y_s}) be the Fourier transform of $(y_t, y_{s,t})$, where $y_{s,t}$ is observed at $t = s, 2s, \dots, n$; $n \mod s$. Then,*

$$\sqrt{s} w_{y_s}(\lambda) = \sum_{j=0}^{s-1} w_y(\lambda + 2\pi j/s) e^{-i(\lambda_k + 2\pi j/s)\cdot s}, \quad \forall \lambda \in \mathscr{B}_s,$$

where $\omega_s = 1/2s$ and $\mathscr{B}_s = [-\omega_s, \omega_s]$.

Notice that Lemma 7 relies on the fact that n is mod s. For n not mod s, the folding process results in a imperfect alignment of the Fourier frequencies, requiring a not straightforward spectral estimation. Furthermore, we can express Lemma 7 in matrix form as $F_s w_{y_s} = FDw_y$, where D, $n \times n$, is a diagonal delay matrix, F, $n \times n$, is a folding matrix, and F_s, $n \times n_s$, is an alignment matrix. See the Appendix A for the functional form of D, F, and F_s.

For finite samples, the HI estimator (3.4) results in a complex $\tilde{c}(j)$. Using $q = [-M+1, \ldots, M]$, as in Hannan (1963b, 1967), we obtain only real values and same asymptotic results of Hidalgo (2000). Thus, given the results of Lemma 7, our proposal for a HI estimator under mixed-frequencies is the following,

$$\tilde{c}(j) = \frac{1}{2M} \sum_{q=-M+1}^{M} \hat{f}^{-1}_{x^d x^d}(\lambda_{2mq}) \hat{f}_{x^d \tilde{y}}(\lambda_{2mq}) e^{ij\lambda_{2mq}}, \quad \lambda_l = 2\pi l/n \tag{3.6}$$

with $j \in \mathbb{Z}$, $w_x^d = Dw_x$ and

$$\underset{(n\times 1)}{w_{\tilde{y}}} = \sqrt{s} \underset{(n \times n_s)}{\begin{pmatrix} I_{n_s} \\ I_{n_s} \\ \vdots \\ I_{n_s} \end{pmatrix}} \underset{(n_s \times 1)}{w_{y_s}}, \tag{3.7}$$

where I_{n_s} is an $n_s \times n_s$ identity matrix. The novelty of this method relies in showing that $\hat{f}_{x^d\tilde{y}}(\lambda) \xrightarrow{p} f_{xy}(\lambda)$, $\lambda \in [0, \pi]$, despite the spectrum of Y_n is folded over itself s times. To demonstrate this claim, notice that the cross-spectrum estimate based on w_x^d and $w_{\tilde{y}}$ for a particular frequency λ, and from Lemma 7, is given by

$$\hat{f}_{x^d\tilde{y}}(\lambda) = \frac{1}{2m+1} \sum_{k=-m}^{m} \left(\sum_{j=0}^{s-1} w_y(\lambda + \lambda_k + 2\pi j/s) e^{-i(\lambda+\lambda_k+2\pi j/s)\cdot s} \right) w_x^d(\lambda + \lambda_k)',$$

assuming, w.l.o.g., that λ lies on $[0, \pi/s)$, we have

$$\begin{aligned} \hat{f}_{x^d\tilde{y}}(\lambda) &= \frac{1}{2m+1} \sum_{k=-m}^{m} w_y(\lambda + \lambda_k) w_x(\lambda + \lambda_k)' \\ &+ \frac{1}{2m+1} \sum_{k=-m}^{m} \left(\sum_{j=1}^{s-1} w_y(\lambda + \lambda_k + 2\pi j/s) e^{-i(\lambda+\lambda_j+2\pi j/s)\cdot s} \right) w_x^d(\lambda + \lambda_k)', \end{aligned}$$

where the second term of RHS, which contains the contribution from aliased frequencies in the aggregation process of y_t, has limit expected value equal to zero due orthogonality of DFT's at different frequencies. This property only holds on strict stationary series. For example for $\lambda \neq 0$, the spectrum of I(1) series is correlated, see Corbae et al. (2002), thus the second term of RHS is not null. Furthermore, at frequency zero the spectrum is unbounded. Then, in Theorem 8 we show that the unfolded cross-spectrum estimate, $\hat{f}_{x^d\tilde{y}}(\lambda)$, is consistent but has a higher variance in comparison with its regular dataset equivalent.

For a lead/lag j, the estimator of covariance matrix $\Omega_{j,lk}$, $l,k = 1,\ldots,p$ is given by,

$$\tilde{\Omega}_{j,lk} = \frac{1}{2M} \sum_{q=-M+1}^{M} \hat{f}_{\tilde{x}\tilde{x}}^{-1}(\lambda_{2mq})\hat{f}_{uu}(\lambda_{2mq})e^{i(l-k)\lambda_{2mq}} \tag{3.8}$$

where $\hat{f}_{uu}(\lambda) = \hat{f}_{\tilde{y}\tilde{y}}(\lambda) - \hat{f}_{\tilde{y}x^d}(\lambda)'\hat{f}_{x^dx^d}^{-1}(\lambda)\hat{f}_{x^d\tilde{y}}(\lambda)$.

4 Asymptotic Theory

Assumption 8. *$\varsigma_t = (u_t, x_t')'$ is a jointly stationary time series with Wold representation $\varsigma_t = \sum_{j=0}^{\infty} A_j\xi_{t-j}$, where $\xi_t = iid(0,\Sigma)$ with finite fourth moments and the coefficients A_j satisfy $\sum_{j=0}^{\infty} j^2||A_j|| < \infty$. The spectral density matrix $f_{\varsigma\varsigma}(\lambda)$ of ς_t satisfies,*

$$f_{\varsigma\varsigma}(\lambda) = \begin{bmatrix} f_{uu}(\lambda) & 0 \\ 0 & f_{xx}(\lambda) \end{bmatrix}$$

with $f_{uu}(\lambda), f_{xx}(\lambda) > 0\ \forall\lambda$.

Assumption 9. *Let $y_t = \sum_{j=-\infty}^{\infty} c(j)x_{t-j} + u_t$ with $Cov(x_t, u_{t-j}) = 0$ for $-\infty < j < \infty$. Let $M = c_\delta n^\delta$, such that $0 < c_\delta < \infty$ and $0 < \delta < 1/3$, with $n^{1/2}\sum_{|j|>M}|c(j)| \to 0$.*

Theorem 8 can be understood as a mixed-frequency version of Theorem 1 of Hidalgo (2000), and Theorem 9 a variant of Theorem 1 of Lütkepohl and Poskitt (1996).

Theorem 8. *Assuming Assumption 1-2, for any finite collection $j_1,\ldots,j_p$, $|j| \leq M$, as $n \to \infty$,*

(i) $n^{1/2}(\tilde{c}(j_1) - c(j_1),\ldots,\tilde{c}(j_p) - c(j_p)) \xrightarrow{d} N(0, \Omega = \{\Omega_{j_rj_l}\}_{r,l=1,\ldots,p})$ where

$$\Omega_{j,lk} = (2\pi)^{-1}\int_{-\pi}^{\pi} f_{xx}^{-1}(\lambda)f_{uu}(\lambda)e^{i(l-k)\lambda}d\lambda \tag{3.9}$$

corresponds to the asymptotic covariance matrix between $\tilde{c}(j_l)$ and $\tilde{c}(j_k)$.

(ii) Equation (3.8) is a consistent estimator of $\Omega_{j,lk}, l,k = 1,\ldots,p$.

Before we state our second theorem, let us define some elements. Similar to Lütkepohl and Poskitt (1996) and Gonçalves and Kilian (2007), let $\ell(M)$ be an arbitrary, $2M+1 \times 1$, vector satisfying $0 < ||\ell(M)||^2 < \infty$, and let $v_M^2 = \ell(M)'\Gamma_M^{-1}B_M\Gamma_M^{-1}\ell(M)$, where $B_M = E(\mathcal{X}_{s,n}^{M\prime}\mathcal{X}_{s,n}^{M}U_{s,n}^2)$ and $\Gamma_M = E(\mathcal{X}_{s,n}^{M\prime}\mathcal{X}_{s,n}^{M})$. With B_M being estimated using a Eicker-White heteroskedasticity-robust covariance matrix estimator, Gonçalves and Kilian (2007).

Theorem 9. *Let Assumption 1-2 be satisfied, then as $M, n \to \infty$,*

$$\ell(M)'\sqrt{n_s - 2M - 1}(\hat{c}_M - c_M)/v_M \xrightarrow{d} N(0,1).$$

with a similar result involving $\tilde{c}$.

In general lines, Theorem 8 is extension of version of Hidalgo (2000)'s Theorem 1 to mixed-frequency datasets. Hidalgo (2000) shows the consistency of the variance and coefficient estimates for stationary long memory series. We, here, restrict the proof only to strict stationary series. Theorem 9 can be related to Theorem 1 of Lütkepohl and Poskitt (1996), where they provide results for an infinite VAR model. In our case, we need to take into consideration the subsampling process on y_t as well the presence of leads.

5 Empirical procedures

A data driven choice of the bandwidth is always preferred over *ad hoc* choices. It has the potential to remove possible bias from the researcher in report only selected results. Hidalgo (2000) suggests two data driven methods for the choice of bandwidth m and/or the number of lags/leads M. The first method mimics the Akaike (1974) AIC criterion and the second is based on the 'leave-one-out' cross-validation method. We follow his Information Criteria proposal, adapted for our case. Let the error $w_{\hat{v}_s}$ be defined as

$$w_{\hat{v}_s}(\lambda) = \sqrt{s}w_{y_s} - \sum_{l=-M}^{M}\sum_{j=0}^{s-1}\tilde{c}(l)w_x(\lambda + 2\pi j/s)e^{i2\pi(s-1+l)}, \quad \lambda \in \mathscr{B}_s.$$

Then, the Information Criteria is given by,

$$AIC_M = \log \sigma_M^2 + 2M/n,$$

with

$$\hat{\sigma}_M^2 = \exp\left(\frac{1}{4\pi M}\sum_{p=-M}^{M}\log \hat{f}_{\tilde{v}\tilde{v}}(\lambda_{2mq})\right),$$

and $w_{\tilde{v}}$ defined as in (3.7).

Hidalgo (2000, 2005) suggest to replace $\hat{f}_{xx}^{-1}(\lambda_0)\hat{f}_{x\tilde{y}}(\lambda_0)$ by $\hat{f}_{xx}^{-1}(\lambda_{2m})\,\hat{f}_{x\tilde{y}}(\lambda_{2m})$ to avoid technical problems with frequency zero estimation. However, his suggestion eliminates an important portion of the series variability when they follow the 'typical spectrum of an Economic variable', see Granger (1966b). Despite our framework does not allow for long memory, the frequency zero can affect finite sample results, since the spectrum at this frequency represent the sample mean. Hidalgo (2000) also suggest that trimming such frequency could introduce a bias problem. Thus, we follow Taufemback (2018b) and avoid problems with frequency zero substituting $w_x(\lambda_0)$ and $w_{y_s}(\lambda_0)$ by $|w_x(\lambda_1)|$ and $|w_{y_s}(\lambda_1)|$, respectively.

For LS method, one could employ any information criteria method, such as AIC or BIC, to obtain a data-driven value of M. In our experiments, we used BIC.

6 Simulation results

As discussed in Taufemback (2018a), macroeconomic series are, generally, reported at levels. First differentiation of an I(1) series, without missing data, give us a desired stationary series. In case of quarterly GDP, we have in levels the sum of last three monthly GDP, say $\bar{y}_t^q = \bar{y}_t^m + \bar{y}_{t-1}^m + \bar{y}_{t-2}^m$; where letters with bars represent variables in levels. For example, first quarter GDP means the sum of January, February, and March. The sum of December, January, and February, if available, would represent $\bar{y}_{t-1}^q$. The difference between $\bar{y}_t^q$ and $\bar{y}_{t-1}^q$ give us the monthly innovation of $\bar{y}_t^q$ series, or its growth rate if we take log differences. However, GDP comes ever quarter with non-overlapping months. Differentiate from a quarter apart bring up some conjunction of shocks in a pattern given by equation (3.10). Assuming, $\bar{y}_t^m = \bar{y}_0^m + \sum_{j=0}^{t-1} y_{t-j}^m$, and Δ_q be the differentiation of two consecutive quarters, then,

$$\begin{aligned}
\Delta_q \bar{y}_{s,t}^q &= \bar{y}_t + \bar{y}_{t-1} + \bar{y}_{t-2} - (\bar{y}_{t-3} + \bar{y}_{t-4} + \bar{y}_{t-5}), \\
&= \sum_{j=0}^{t} y_{t-j}^m + \sum_{j=0}^{t-1} y_{t-j}^m + \sum_{j=0}^{t-2} y_{t-j}^m - \sum_{j=0}^{t-3} y_{t-j}^m - \sum_{j=0}^{t-4} y_{t-j}^m - \sum_{j=0}^{t-5} y_{t-j}^m \qquad (3.10) \\
&= y_t^m + 2y_{t-1}^m + 3y_{t-2}^m + 2y_{t-3}^m + y_{t-4}^m.
\end{aligned}$$

Consequently, to regress $\Delta_q \bar{y}_{s,t}$ on monthly indicators, x_t, we need to reproduce the same pattern of aggregation on monthly series, say $\Delta_q \bar{x}_t = x_t + 2x_{t-1} + 3x_{t-2} + 2x_{t-3} + x_{t-4}$. Alternatively, one can generate a new series representing the sum of last three month at levels and then differentiate.

To test joint causality of an entire series on other, we propose the following data generated process,

$$DGP.1 \quad \begin{array}{rcl} y_t & = & 0.5y_{t-1} + 0.6x_{t-1} + 0.3x_{t-2} + u_{1,t} - 0.4u_{1,t-1}, \\ z_t & = & 0.6z_{t-1} + u_{2,t} - 0.3u_{2,t-1}, \end{array} , u_t \sim N(0,1),$$

where x_t is collected in two fashions: $x_t = z_t$ and $x_t = z_t + 2z_{t-1} + 3z_{t-2} + 2z_{t-3} + z_{t-4}$, respectively. And y_t is sub-sampled in the following schemes: quarterly-monthly, annually-quarterly, annually-monthly, and monthly-daily. The correspondent values of s are $\{3, 4, 12, 20\}$. For the monthly-daily scenario, we follow Götz et al. (2016) suggestion of $s = 20$. They argue that with real data 20 is the maximum number of working days available. For those months with more than 20 working days, they remove the first days of the month sample.

We also propose a variant of DGP.1, called DGP.1x, where the subsample variable is x_t rather than y_t. We provide results for the same set of subsample ratio described previously. The sample size was chosen to be $\{50, 100, 150\}$ for the low-frequency variable, with exception of $\{50, 75, 100\}$ for $s = 20$.

Notice that DGP.1, and consequentially DGP.1x, admits an infinite order representation concerning lags due to the MA(1) term. Under this formulation, X_n causes $Y_{s,n}$, but with no inverse causality, i.e. $Y_{s,n} \not\rightarrow X_n$. We expect a rejection rate of the null hypothesis close to the unit in the first case and close to the nominal rate, 5%, in the second. Moreover, since we have the opposite causality relation for DGP.1x, we also expect the opposite results regarding the rejection rate of the null.

Unreported simulation results involving x_t such that it follows the simulated GDP sampling scheme are very unsatisfactory for both MF-HI and LS estimation. The reason behind this behavior is that, due the aggregation, the low frequencies spectrum explain the majority of series variability. For reduced sample sizes, spectral methods, such MF-HI estimation, do not perform well under these conditions. The same can be inferred for least squares estimation since Lütkepohl and Poskitt (1996) faced similar problems. Using an MA(1) model, their findings suggest that as close the MA coefficient to the unit, less reliable the tests appears. However, this problem can be contoured using the unaggregated variable z_t as input, since with enough lags the aggregation pattern can be reproduced. In practice, we observe z_t rather than x_t. For example, using monthly macroeconomic variables, it is necessary to aggregate them every three months to match the aggregation pattern of a quarterly variable. Thus, using z_t we avoid an extra data preparation step.

Thus, we only report here results for simulated GDP sampling scheme with unaggregated dependent value. Table C.3-1 and C.3-3 shows that MF-HI estimator reported better size control when compared with the least squares estimation, Table C.3-2 and C.3-4. The MF-HI method reported a slight lower rejection rate when causality is present. The least squares showed a convergence to the nominal rejection of $y_{s,t} \nrightarrow x_t$ as the number of observations increased, but an over rejection for all cases whenever $x_{s,t} \nrightarrow y_t$.

Notice that the result analysis may be not straightforward for the distinct values of s. In a quarterly-monthly scenario, with 50 quarterlies we have 150 monthly observations, but in a monthly-daily, with 50 months of data, we have around 1,000 daily observations. Thus, the expected drop in efficiency caused by a higher value of s is compensated by the high amount of data in the dependent value.

With DGP.2 we aim to study the rejection under the null and the alternative for lead and lags. Here, we do not expect that MF-HI estimator to perform well. Its inefficiency is well documented for regular datasets, see Cargill and Meyer (1974) for a comparison between least squares and MF-HI estimator. Thus, with mixed-frequency datasets, we expect an even worst performance.

To measure causality at different horizons, let us introduce our second experiment,

$$DGP.2 \quad \begin{array}{rcl} y_t &=& \sum_{j=1}^{6} \beta_{j-1} x_{t-j} + 0.5 u_{1,t}, \quad \beta_j = 0.9^j, \\ z_t &=& 0.5 z_{t-1} + u_{2,t}, \end{array} , u_t \sim N(0,1),$$

again x_t is collected in two fashions: first as skip sampling stock variables and second as a simulation of the GDP sampling scheme, i.e., $x_t = z_t$ and $x_t = z_t + 2z_{t-1} + 3z_{t-2} + 2z_{t-3} + z_{t-4}$, respectively. Series y_t is sub-sampled in the following schemes: quarterly-monthly, annually-quarterly, annually-monthly, and monthly-daily. The correspondent values of s are $\{3, 4, 12, 20\}$.

Here, we aim two objectives. The first is study the rejection rate under the null whether an specific lag (lead) of x_t causes (do not cause) $y_{s,t}$. We expect a high rejection concerning lag causality and a nominal rejection rate, 5%, for lead causality. Our second objective is to compare how both methods perform concerning finite sample bias. As mention before, MF-HI method is not an accurate method but interest to study its behavior compared with a more efficient method such the LS.

Table C.3-5 and C.3-6, in Appendix C, presents the results for stock variables lags. In neither cases, the estimated coefficients present any significative finite sample bias. However, the rejection rate of MF-HI is inferior to those using LS estimator. MF-HI method reports a satisfactory rejection rate but not as precise as the LS approach.

Again, simulation results under the GDP sampling scheme with the dependent value are far from acceptable, presenting large finite sample bias, especially for the MF-HI estimator. Increasing the number of observations and/or reducing the values of β has a positive, but small, impact on the results. Thus, we again regress the aggregated $y_{s,t}$ on z_t to obtain better results. However, using z_t, instead of x_t, results in estimated coefficients which are now an aggregation of β_j, for example let $\tilde{\beta}$ be the estimated coefficient of $y_{s,t}$ on z_t, then

$$\begin{aligned}
\tilde{\beta}_0 &= \beta_0, & \tilde{\beta}_5 &= \beta_1 + 2\beta_2 + 3\beta_3 + 2\beta_4 + \beta_5,\\
\tilde{\beta}_1 &= 2\beta_0 + \beta_1, & \tilde{\beta}_6 &= \beta_2 + 2\beta_3 + 3\beta_4 + 2\beta_5,\\
\tilde{\beta}_2 &= 3\beta_0 + 2\beta_1 + \beta_2, & \tilde{\beta}_7 &= \beta_3 + 2\beta_4 + 3\beta_5,\\
\tilde{\beta}_3 &= 2\beta_0 + 3\beta_1 + 2\beta_2 + \beta_3, & \tilde{\beta}_8 &= \beta_4 + 2\beta_5,\\
\tilde{\beta}_4 &= \beta_0 + 2\beta_1 + 3\beta_2 + 2\beta_3 + \beta_4, & \tilde{\beta}_9 &= \beta_5.
\end{aligned}$$

For MF-HI estimation, the results for the estimation of $\tilde{\beta}$ are inferior to those with stock variables but with no finite sample bias, see Table C.3-7. LS estimation shows a superior performance for test causality at different horizons, Table C.3-3. In terms of leads, Tables C.3-9 and C.3-10 reports results for stock variables and Tables C.3-11 and C.3-12 under simulated GDP sampling scheme, under the null. Both methods, with both sampling schemes, present good results approaching the nominal rejection rate of 5%. The LS method presents overall better size and lower average bias, see Table C.3-3.

Summarizing, the MF-HI method shows better results testing overall causality from one series to another, for both directions of causality. However, the LS method proved to be the more indicated in test causality of one specific lag to the dependent variable.

7 Empirical results

Here, as in Ghysels et al. (2016), we study the prediction power between a list of US macroeconomic indicators and the US GDP, see Table 3-1 for the used series in this study. Ghysels et al. (2016) studied the causality between US quarterly GDP and two monthly variables, Oil prices and Consumer price index. Here, we cluster our variable set in two groups, in the first we have Employees payrolls, Personal income, Industrial production, and Sales. In the second group, we have S&P index, Crude Oil, and Consumer Price Index. We test causality from these two group to GDP and from GDP to these groups.

Ghysels et al. (2016) compare their proposed mixed-frequency causality test with using standard methods after performing temporal aggregation on high-frequency variables. The distinct methods result in also distinct outcomes. Under mixed-frequencies the results of Ghysels et al. (2016) implies that OIL causes CPI at $j = 1, 4$, CPI causes GDP at $j = 3$, and GDP causes CPI at $j = 1$. For the temporal aggregation method, CPI causes OIL at $j = 1$ and OIL causes GDP at $j = 2, 4$.

Table 3-1: US indicators from January of 1959 to June of 2015.

	Description
	Quarterly
GDP	Real GDP[1] (billions of chained 2009 dollars, SA, AR).
	Monthly
EMP	Employees on non-agricultural payrolls[2] (SA);
INC	Personal income less transfer payments[3] (2009 dollars, SA, AR);
IIP	Index of industrial production[4] (2012=100, SA);
SLS	Manufacturing and trade sales[3] (2009 dollars, SA).
S&P	S&P's Common Stock Price Index: Composite[5]
OIL	Crude Oil, spliced WTI and Cushing[5]
CPI	Consumer Price Index : All Items[5]

Notes: SA means 'seasonally adjusted' and AR means 'annual rate'. Source: (1) BEA - Bureau of Economic Analysis, U.S. Department of Commerce, (2) Bureau of Labor Statistics, U.S. Department of Labor, (3) The Conference Board, (4) Federal Reserve, United States, (5) St. Louis FED.

Let us, arbitrarily, define $|j| \leq 5$ as short run and $|j| > 5$ long run. Then, combining the results of our both methods, Table 3-2 and 3-3, we found causality to GDP in the short-run from EMP, and SLS. Also, causality in the long-run from EMP, IIP, and CPI. From causality of the low-frequency variable, GDP, to high-frequency variables, we found causality in the short run for EMP, INC, and in the long run for IIP, SLS, and CPI. Figures 3-1 to 3-4 report graphically the leads/lags coefficients and their correspondent 95% confidence interval. Both methods only agree with joint causality for SLS causing GDP.

Our results are not in consonance with Ghysels et al. (2016). Using mixed-frequency, they found short-run causality from GDP to CPI, while we found long-run causality. The same with CPI to GDP. However, our two methods present similar results, which reinforce our conclusions.

8 Conclusion

We propose two novel nonparametric causality test for mixed-frequency datasets. Under a two-side unrestricted distributed lag model, we proposed a Least Squares and a Hannan-Inefficient estimator approaches.

We assume that the low-frequency and the high-frequency variables are generated at the same frequency, but we systematic miss some observations for the LF variable. Causality tests based on mixed-frequency VAR, Ghysels et al. (2016), and Götz et al. (2016), assume a mismatch in the generation process, which implies that models for regular datasets cannot be employed. Furthermore, in contrary to their methods, our proposal test does not suffer from parameter proliferation.

Simulations show that our two proposed methods complement each other. Where MF-HI is more indicated for joint causality test, LS method is better in measuring causality at a given horizon. Furthermore, our empirical application finds causality at different lags and leads for GDP and US monthly indicators.

Table 3-2: MF-HI method multivariate causality tests for GDP, quarterly, and US macroeconomic indicators, monthly.

$\mathbb{H}_0^j(X_n \nrightarrow Y_n)$

j	EMP	INC	IIP	SLS	S&P	OIL	CPI
1	0.211	0.701	0.679	**0.044**	0.995	0.863	0.345
2	**0.036**	0.310	0.916	**0.003**	**0.027**	0.178	0.479
3	0.929	0.749	0.098	**0.003**	0.299	0.594	0.113
4	0.726	0.961	0.431	0.868	0.945	0.977	0.430
5	0.588	0.387	**0.012**	0.729	**0.041**	0.464	0.568
6	0.630	0.705	0.718	0.052	0.054	0.619	**0.036**
7	0.533	0.344	0.655	0.127	0.528	0.549	0.375
8	**0.037**	0.584	**0.037**	0.281	0.066	0.602	0.753
9	0.427	0.798	0.452	0.358	0.072	0.208	**0.029**
10	0.088	0.284	0.881	0.899	0.614	0.336	0.190

$\mathbb{H}_0(X_n^p \nrightarrow Y_n)$

	EMP	INC	IIP	SLS	S&P	OIL	CPI
	0.320	0.719	0.468	**0.029**	0.097	0.982	0.145

$\mathbb{H}_0^j(Y_n \nrightarrow X_n)$

j	EMP	INC	IIP	SLS	S&P	OIL	CPI
-1	0.704	**0.004**	0.913	0.358	0.780	0.498	0.673
-2	**0.019**	**0.018**	0.348	0.535	0.881	0.951	0.354
-3	0.799	0.155	0.869	0.937	0.722	0.142	0.940
-4	0.258	0.065	0.066	0.660	0.098	0.514	0.440
-5	0.354	0.323	0.515	0.970	0.267	0.466	0.973
-6	0.977	0.500	**0.014**	**0.016**	0.375	0.064	0.554
-7	0.163	0.703	0.088	0.954	0.316	0.762	0.057
-8	0.419	0.740	0.201	0.216	0.447	0.169	0.861
-9	0.813	0.875	0.330	0.367	0.948	0.919	**0.001**
-10	0.074	0.401	0.763	0.902	0.726	0.909	0.855

$\mathbb{H}_0(Y_n \nrightarrow X_n^p)$

	EMP	INC	IIP	SLS	S&P	OIL	CPI
	0.509	0.366	0.168	0.433	0.682	0.636	0.073

Notes: Number of lags choose by AIC_M, see Section 5, with $M \in [2n^{1/3}, 3n^{1/3}]$. We only report 10 lead/lags. Bold values represent p-values below 5%.

Table 3-3: LS method multivariate causality tests for GDP, quarterly, and US macroeconomic indicators, monthly.

$\mathbb{H}_0^j(X_n \nrightarrow Y_n)$

j	EMP	INC	IIP	SLS	S&P	OIL	CPI
1	0.118	0.619	0.972	**0.012**	0.694	0.889	0.995
2	**0.021**	0.262	0.249	**0.040**	0.300	0.963	0.693
3	0.830	0.221	0.783	0.537	0.998	0.644	0.911
4	0.341	0.151	0.199	0.316	0.252	0.064	0.429
5	0.591	**0.043**	0.467	0.769	0.143	**0.007**	0.777
6	0.801	0.224	0.059	0.421	0.631	0.787	0.764
7	0.557	0.089	0.299	0.336	0.321	0.722	**0.013**
8	0.991	0.900	**0.023**	0.476	0.473	0.590	0.695
9	0.698	0.911	0.760	0.878	0.186	0.957	0.626
10	**0.031**	0.906	0.590	0.075	0.941	0.094	0.394

$\mathbb{H}_0(X_n^p \nrightarrow Y_n)$

EMP	INC	IIP	SLS	S&P	OIL	CPI
0.226	0.731	0.078	**0.041**	**0.007**	0.118	0.171

$\mathbb{H}_0^j(Y_n \nrightarrow X_n)$

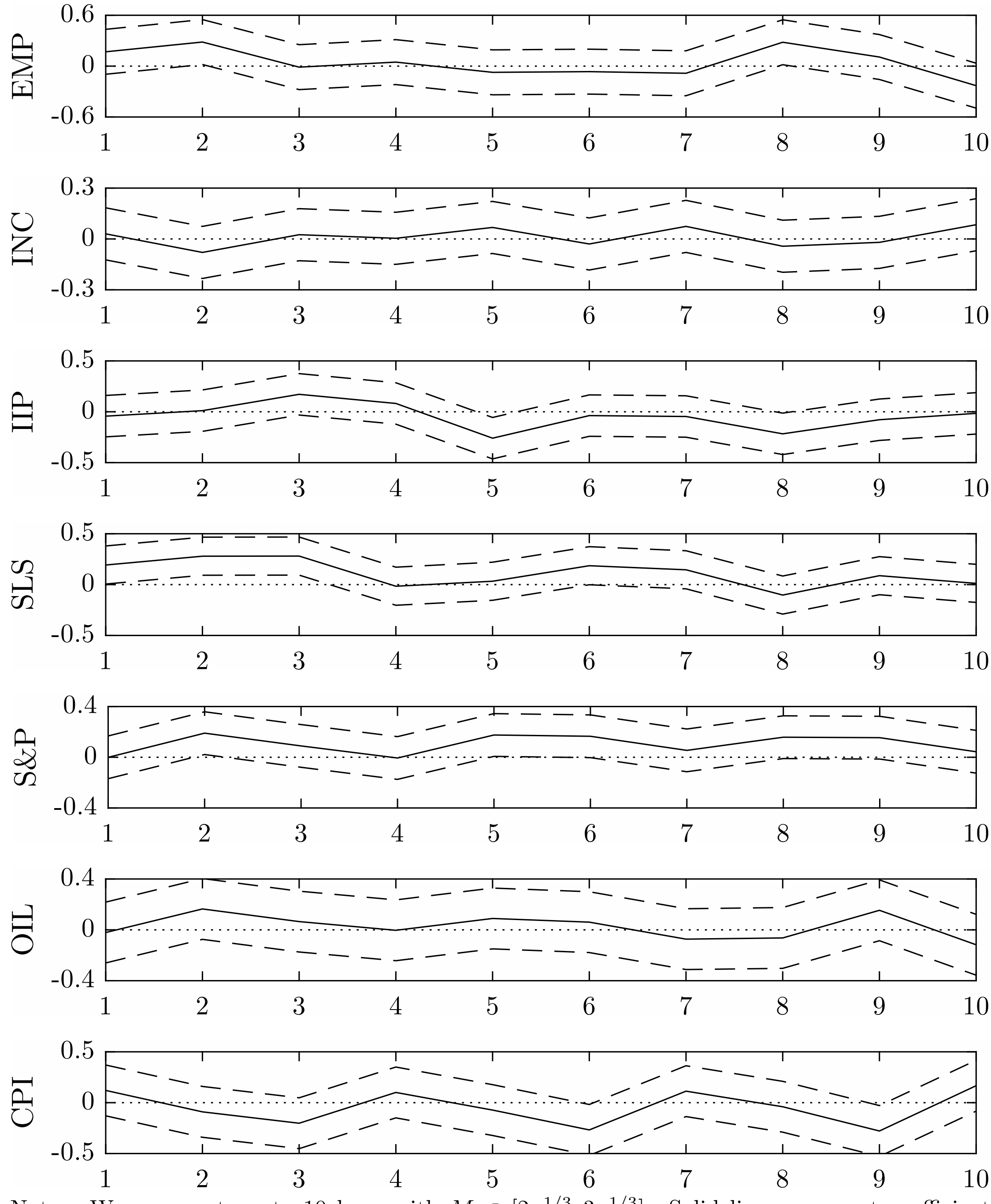
j	EMP	INC	IIP	SLS	S&P	OIL	CPI
-1	0.105	0.309	0.326	0.325	0.101	0.445	0.206
-2	**0.034**	**0.050**	0.292	0.556	0.570	0.787	0.141
-3	0.253	0.559	1.000	0.330	0.140	0.671	0.098
-4	0.159	**0.030**	0.133	0.927	0.395	**0.012**	0.438
-5	**0.041**	0.134	0.968	0.852	0.906	0.221	**0.011**
-6	0.451	0.727	**0.004**	**0.002**	0.743	0.750	0.627
-7	0.686	0.828	0.088	**0.049**	0.440	0.189	0.752
-8	0.878	0.702	**0.009**	0.129	0.148	0.728	**0.036**
-9	0.579	0.275	0.365	0.190	0.670	0.890	0.326
-10	0.270	0.352	0.957	0.626	0.102	0.780	0.697

$\mathbb{H}_0(Y_n \nrightarrow X_n^p)$

EMP	INC	IIP	SLS	S&P	OIL	CPI
0.127	**0.039**	**0.014**	**0.028**	**0.043**	0.562	**0.029**

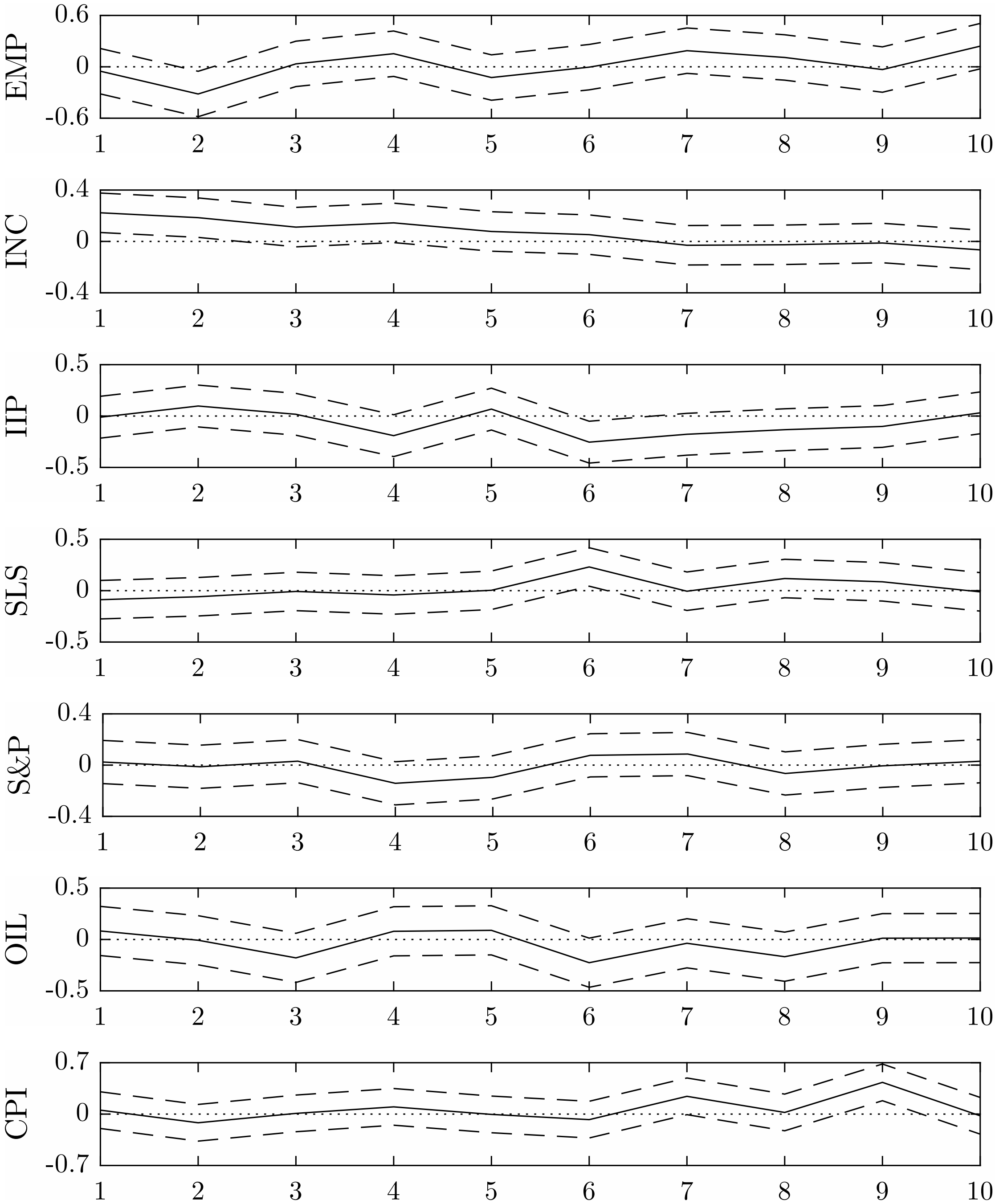
Notes: Number of lags choose by BIC, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$. We only report 10 lead/lags. Bold values represent p-values below 5%.

Figure 3-1: Lag coefficients and 95% CI for MF-HI method multivariate causality tests for GDP, quarterly, and US macroeconomic indicators, monthly.



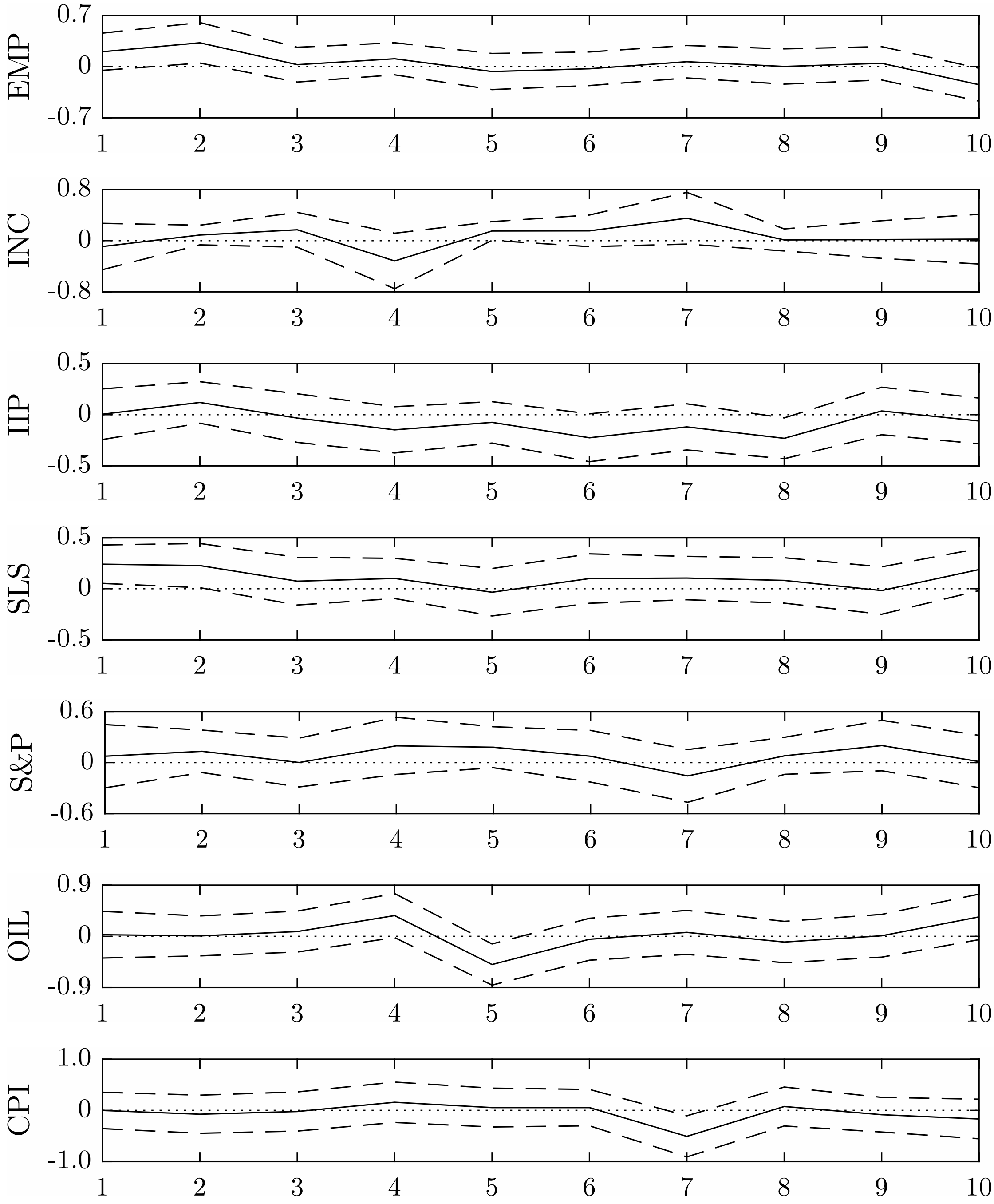
Notes: We represent up to 10 lags, with $M \in [2n^{1/3}, 3n^{1/3}]$. Solid lines represent coefficient values and dashed lines the 95% Confidence interval.

Figure 3-2: Lead coefficients and 95% CI for MF-HI method multivariate causality tests for GDP, quarterly, and US macroeconomic indicators, monthly.



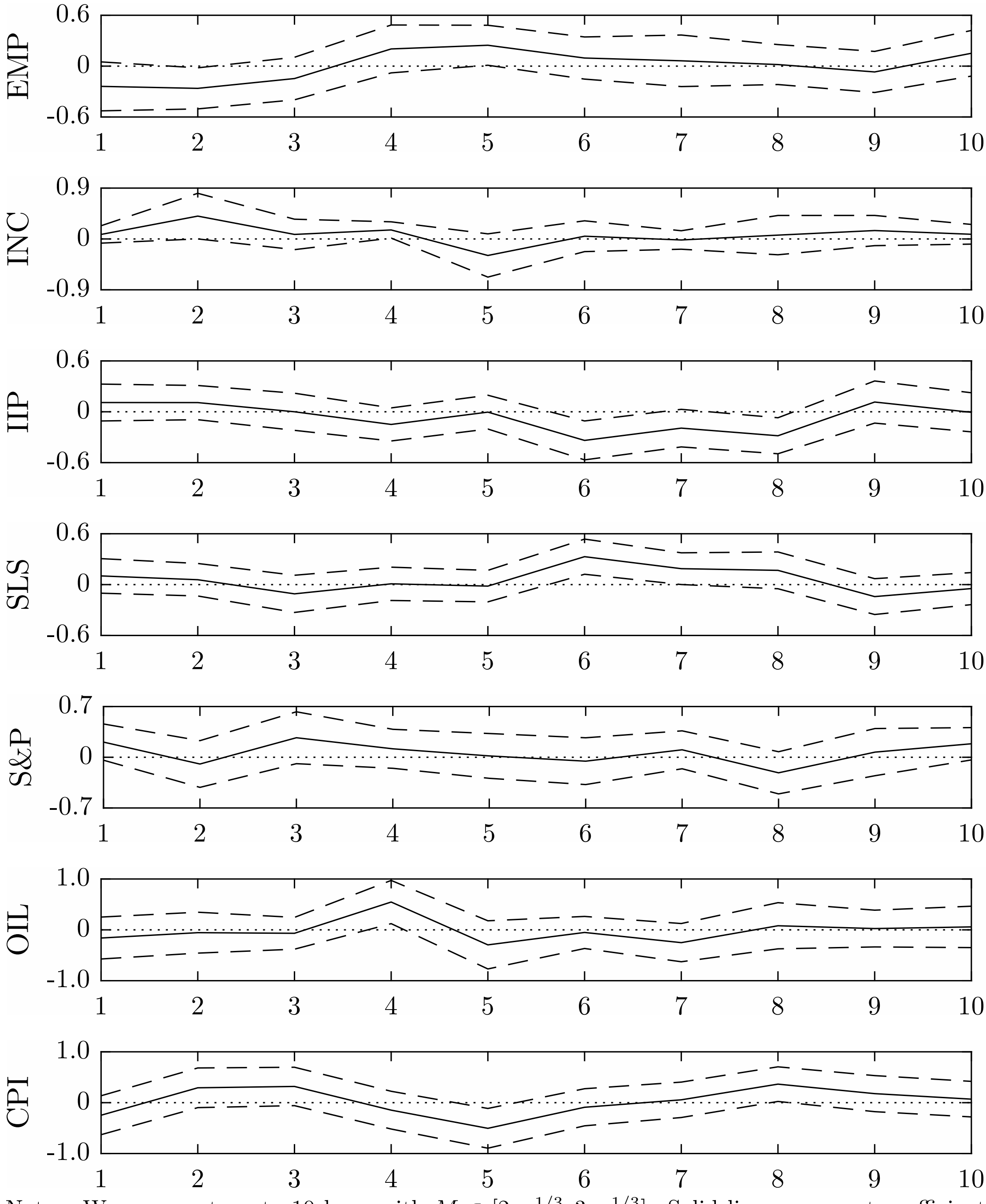
Notes: We represent up to 10 lags, with $M \in [2n^{1/3}, 3n^{1/3}]$. Solid lines represent coefficient values and dashed lines the 95% Confidence interval.

Figure 3-3: Lag coefficients and 95% CI for LS method multivariate causality tests for GDP, quarterly, and US macroeconomic indicators, monthly.



Notes: We represent up to 10 lags, with $M \in [2ns^{1/3}, 3ns^{1/3}]$. Solid lines represent coefficient values and dashed lines the 95% Confidence interval.

Figure 3-4: Lead coefficients and 95% CI for LS method multivariate causality tests for GDP, quarterly, and US macroeconomic indicators, monthly.



Notes: We represent up to 10 lags, with $M \in [2ns^{1/3}, 3ns^{1/3}]$. Solid lines represent coefficient values and dashed lines the 95% Confidence interval.

Appendix 3.A

The functional form of D_l, F, and F_s is given by: D, $n \times n$, is a diagonal matrix where $d_{kk} = e^{i\lambda_k(s-1l)}$, for $\lambda_k = 2\pi k/n$ with $k = 0, \ldots, n-1$ and

$$F = \left[\begin{array}{c:c:c:c} IO_\tau & IO_\tau & \ldots & IO_\tau \\ \hdashline \multicolumn{4}{c}{O_{n-n_s \times n}} \\ \hdashline IO_{\tau^*} & \ldots & IO_{\tau^*} & IO_{\tau^*} \end{array}\right]_{n \times n}, \quad F_s = \sqrt{s}\left[\begin{array}{c:c} I_{\tau\times\tau} & O_{\tau \times n_s-\tau} \\ \hdashline \multicolumn{2}{c}{O_{n-n_s \times n_s}} \\ \hdashline O_{\tau^* \times n_s-\tau^*} & I_{\tau^*\times\tau^*} \end{array}\right]_{n \times n_s},$$

where $IO_\tau = [I_{\tau\times\tau} \quad O_{\tau\times\tau^*}]$, $IO_{\tau^*} = [O_{\tau^*\times\tau} \quad I_{\tau^*\times\tau^*}]$, $\tau = \lfloor n_s/2 \rfloor + 1$ and $\tau^* = n_s - \tau$.

Appendix 3.B

Proofs...

Appendix 3.C

Table C.3-1: Simulated size and power using MF-HI method, for DGP.1 and DGP.1x under systematic skip sampling stock variables.

	n	DGP.1 $\mathbb{H}_1 : X_n \to Y_{s,n}$	DGP.1 $\mathbb{H}_0 : Y_{n,s} \not\to X_n$	DGP.1x $\mathbb{H}_0 : X_n \not\to Y_n$	DGP.1x $\mathbb{H}_1 : Y_n \to X_n$
$s=3$	50	0.998	0.090	0.104	0.974
	100	1.000	0.054	0.067	1.000
	150	1.000	0.050	0.058	1.000
$s=4$	50	0.997	0.071	0.077	0.954
	100	1.000	0.047	0.052	1.000
	150	1.000	0.046	0.061	1.000
$s=12$	50	0.963	0.039	0.041	0.847
	100	1.000	0.031	0.031	0.994
	150	1.000	0.028	0.034	1.000
$s=20$	50	0.941	0.034	0.027	0.781
	75	1.000	0.024	0.024	0.990
	100	1.000	0.018	0.038	0.999

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n^{1/3}, 3n^{1/3}]$ and 50,000 MC repetitions.

Table C.3-2: Simulated size and power using LS method, for DGP.1 and DGP.1x under systematic skip sampling stock variables.

		DGP.1		DGP.1x	
	n	$\mathbb{H}_1 : X_n \to Y_{s,n}$	$\mathbb{H}_0 : Y_{n,s} \nrightarrow X_n$	$\mathbb{H}_0 : X_n \nrightarrow Y_n$	$\mathbb{H}_1 : Y_n \to X_n$
$s = 3$	50	0.997	0.157	0.157	0.973
	100	1.000	0.082	0.083	1.000
	150	1.000	0.065	0.090	1.000
$s = 4$	50	0.998	0.143	0.152	0.982
	100	1.000	0.078	0.099	1.000
	150	1.000	0.072	0.094	1.000
$s = 12$	50	1.000	0.134	0.138	0.987
	100	1.000	0.083	0.093	1.000
	150	1.000	0.069	0.085	1.000
$s = 20$	50	0.999	0.128	0.139	0.987
	75	1.000	0.070	0.094	1.000
	100	1.000	0.066	0.082	1.000

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. The number of leads/lags was chosen according to AIC method, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$ and 50,000 MC repetitions.

Table C.3-3: Simulated size and power using MF-HI method, for DGP.1 and DGP.1x under simulated GDP sampling scheme.

		DGP.1		DGP.1x	
	n	$\mathbb{H}_1 : X_n \to Y_{s,n}$	$\mathbb{H}_0 : Y_{n,s} \not\to X_n$	$\mathbb{H}_0 : X_n \not\to Y_n$	$\mathbb{H}_1 : Y_n \to X_n$
$s = 3$	50	1.000	0.099	0.096	1.000
	100	1.000	0.049	0.073	1.000
	150	1.000	0.034	0.062	1.000
$s = 4$	50	1.000	0.072	0.085	0.999
	100	1.000	0.038	0.066	1.000
	150	1.000	0.033	0.061	1.000
$s = 12$	50	1.000	0.038	0.065	0.982
	100	1.000	0.031	0.056	1.000
	150	1.000	0.032	0.047	1.000
$s = 20$	50	1.000	0.045	0.071	0.974
	75	1.000	0.022	0.043	1.000
	100	1.000	0.015	0.046	1.000

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n^{1/3}, 3n^{1/3}]$ and 5,000 MC repetitions.

Table C.3-4: Simulated size and power using LS method, for DGP.1 and DGP.1x under simulated GDP sampling scheme.

	n	DGP.1 $\mathbb{H}_1 : X_n \to Y_{s,n}$	DGP.1 $\mathbb{H}_0 : Y_{n,s} \not\to X_n$	DGP.1x $\mathbb{H}_0 : X_n \not\to Y_n$	DGP.1x $\mathbb{H}_1 : Y_n \to X_n$
$s = 3$	50	1.000	0.178	0.197	1.000
	100	1.000	0.086	0.159	1.000
	150	1.000	0.074	0.176	1.000
$s = 4$	50	1.000	0.177	0.190	1.000
	100	1.000	0.090	0.156	1.000
	150	1.000	0.070	0.186	1.000
$s = 12$	50	1.000	0.159	0.164	1.000
	100	1.000	0.076	0.166	1.000
	150	1.000	0.062	0.190	1.000
$s = 20$	50	1.000	0.148	0.179	1.000
	75	1.000	0.074	0.162	1.000
	100	1.000	0.066	0.191	1.000

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. The number of leads/lags was chosen according to AIC method, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$ and 50,000 MC repetitions.

Table C.3-5: Average lag coefficient estimation and the rejection rate of $\mathbb{H}_0^j$: $c(j) = 0, j > 0$ for MF-HI method and DGP.2 under systematic skip sampling stock variables.

	n	$\tilde{c}(1)$	$\tilde{c}(2)$	$\tilde{c}(3)$	$\tilde{c}(4)$	$\tilde{c}(5)$	$\tilde{c}(6)$
$s = 3$	50	1.013	0.911	0.767	0.688	0.591	0.460
		(0.511)	(0.447)	(0.343)	(0.300)	(0.230)	(0.155)
	100	1.005	0.908	0.789	0.683	0.616	0.501
		(0.771)	(0.705)	(0.597)	(0.489)	(0.418)	(0.297)
	150	0.988	0.870	0.776	0.681	0.606	0.508
		(0.901)	(0.826)	(0.753)	(0.641)	(0.551)	(0.422)
$s = 4$	50	1.013	0.900	0.832	0.659	0.589	0.490
		(0.526)	(0.465)	(0.414)	(0.305)	(0.246)	(0.183)
	100	0.992	0.890	0.812	0.668	0.602	0.522
		(0.744)	(0.652)	(0.589)	(0.447)	(0.378)	(0.299)
	150	1.010	0.910	0.823	0.714	0.638	0.556
		(0.922)	(0.866)	(0.798)	(0.688)	(0.599)	(0.485)
$s = 12$	50	1.009	0.905	0.793	0.718	0.639	0.571
		(0.499)	(0.422)	(0.353)	(0.298)	(0.254)	(0.218)
	100	1.000	0.871	0.796	0.696	0.644	0.540
		(0.740)	(0.625)	(0.555)	(0.451)	(0.413)	(0.305)
	150	1.019	0.909	0.807	0.730	0.637	0.581
		(0.907)	(0.829)	(0.744)	(0.653)	(0.549)	(0.485)
$s = 20$	50	1.036	1.021	0.885	0.759	0.602	0.483
		(0.077)	(0.067)	(0.054)	(0.044)	(0.029)	(0.019)
	75	1.072	0.935	0.831	0.570	0.614	0.365
		(0.129)	(0.111)	(0.087)	(0.070)	(0.059)	(0.034)
	100	1.025	0.891	0.791	0.681	0.533	0.386
		(0.143)	(0.123)	(0.098)	(0.081)	(0.058)	(0.043)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 0.90, 0.81, 0.73, 0.66\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n^{1/3}, 3n^{1/3}]$ and 5,000 MC repetitions.

Table C.3-6: Average lag coefficient estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j > 0$ for LS method and DGP.2 under systematic skip sampling stock variables.

	n	$\hat{c}(1)$	$\hat{c}(2)$	$\hat{c}(3)$	$\hat{c}(4)$	$\hat{c}(5)$	$\hat{c}(6)$
$s = 3$	50	1.003	0.900	0.811	0.729	0.655	0.589
		(1.000)	(1.000)	(1.000)	(1.000)	(0.999)	(0.998)
	100	0.998	0.900	0.810	0.729	0.657	0.592
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	1.000	0.900	0.810	0.728	0.657	0.589
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
$s = 4$	50	1.002	0.897	0.811	0.729	0.657	0.590
		(1.000)	(1.000)	(1.000)	(0.999)	(0.999)	(0.998)
	100	0.999	0.902	0.808	0.729	0.657	0.590
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	1.000	0.900	0.809	0.730	0.656	0.590
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
$s = 12$	50	0.999	0.899	0.811	0.726	0.658	0.590
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	100	1.000	0.899	0.810	0.729	0.657	0.590
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	1.000	0.899	0.810	0.728	0.658	0.590
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
$s = 20$	50	1.001	0.903	0.808	0.728	0.655	0.590
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(0.999)
	100	0.998	0.900	0.811	0.730	0.656	0.589
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	1.001	0.899	0.809	0.731	0.656	0.591
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 0.90, 0.81, 0.73, 0.66\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$ and 5,000 MC repetitions.

Table C.3-7: Average lag coefficient estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j > 0$ for MF-HI method and DGP.2 under simulated GDP sampling scheme. .

	n	$\tilde{c}(1)$	$\tilde{c}(2)$	$\tilde{c}(3)$	$\tilde{c}(4)$	$\tilde{c}(5)$	$\tilde{c}(6)$
$s = 3$	50	0.987	3.235	5.505	6.595	6.848	5.517
		(0.120)	(0.194)	(0.340)	(0.421)	(0.439)	(0.330)
	100	1.110	3.036	5.440	6.811	6.966	5.803
		(0.091)	(0.196)	(0.459)	(0.637)	(0.648)	(0.509)
	150	1.057	2.875	5.380	6.679	6.802	5.743
		(0.092)	(0.242)	(0.605)	(0.784)	(0.799)	(0.658)
$s = 4$	50	1.136	3.086	5.708	6.504	6.837	5.718
		(0.105)	(0.160)	(0.325)	(0.383)	(0.410)	(0.311)
	100	1.092	2.887	5.678	6.666	6.813	5.903
		(0.098)	(0.189)	(0.458)	(0.574)	(0.591)	(0.480)
	150	1.092	2.928	5.798	6.946	7.233	6.304
		(0.104)	(0.270)	(0.676)	(0.806)	(0.844)	(0.745)
$s = 12$	50	1.050	2.962	5.733	6.908	7.150	6.483
		(0.084)	(0.142)	(0.285)	(0.363)	(0.401)	(0.333)
	100	0.972	2.919	5.535	6.873	7.118	6.320
		(0.076)	(0.175)	(0.423)	(0.587)	(0.618)	(0.515)
	150	1.000	2.974	5.695	7.070	7.277	6.452
		(0.077)	(0.235)	(0.605)	(0.786)	(0.815)	(0.722)
$s = 20$	50	1.569	3.180	6.168	6.748	7.201	5.359
		(0.065)	(0.063)	(0.059)	(0.045)	(0.029)	(0.021)
	100	1.497	3.637	5.915	6.843	5.935	4.547
		(0.104)	(0.097)	(0.090)	(0.079)	(0.050)	(0.042)
	150	1.145	3.506	4.992	6.750	6.511	4.827
		(0.094)	(0.097)	(0.084)	(0.091)	(0.074)	(0.057)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 2.90, 5.60, 7.05, 7.34, 6.61\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n^{1/3}, 3n^{1/3}]$ and 5,000 MC repetitions.

Table C.3-8: Average lag coefficient estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j > 0$ for LS method and DGP.2 under simulated GDP sampling scheme.

	n	$\hat{c}(1)$	$\hat{c}(2)$	$\hat{c}(3)$	$\hat{c}(4)$	$\hat{c}(5)$	$\hat{c}(6)$
$s = 3$	50	1.001	2.899	5.611	7.048	7.346	6.614
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	100	1.001	2.901	5.610	7.049	7.345	6.608
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	0.999	2.900	5.609	7.049	7.343	6.609
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
$s = 4$	50	1.000	2.900	5.611	7.050	7.345	6.608
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	100	0.999	2.900	5.611	7.048	7.344	6.610
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	0.999	2.901	5.610	7.048	7.344	6.610
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
$s = 12$	50	0.997	2.903	5.609	7.049	7.341	6.612
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	100	0.999	2.902	5.608	7.049	7.345	6.610
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	1.001	2.899	5.611	7.049	7.343	6.610
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
$s = 20$	50	1.002	2.898	5.610	7.050	7.344	6.609
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	100	1.002	2.900	5.610	7.050	7.342	6.610
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
	150	1.000	2.901	5.610	7.051	7.342	6.610
		(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 2.90, 5.60, 7.05, 7.34, 6.61\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$ and 5,000 MC repetitions.

Table C.3-9: Average lead coefficient bias estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j < 0$ for MF-HI method and DGP.2 under systematic skip sampling stock variables.

	n	$\tilde{c}(-1)$	$\tilde{c}(-2)$	$\tilde{c}(-3)$	$\tilde{c}(-4)$	$\tilde{c}(-5)$	$\tilde{c}(-6)$
$s = 3$	50	-0.003	0.009	-0.008	0.007	0.005	-0.004
		(0.068)	(0.064)	(0.056)	(0.054)	(0.049)	(0.041)
	100	0.000	0.003	-0.001	0.004	-0.003	-0.001
		(0.073)	(0.073)	(0.075)	(0.064)	(0.067)	(0.064)
	150	0.011	-0.012	0.016	-0.011	0.005	-0.004
		(0.064)	(0.065)	(0.062)	(0.064)	(0.061)	(0.053)
$s = 4$	50	0.009	-0.006	-0.005	0.002	0.013	-0.004
		(0.069)	(0.066)	(0.065)	(0.062)	(0.057)	(0.048)
	100	0.004	-0.016	0.010	0.003	0.002	-0.001
		(0.075)	(0.081)	(0.084)	(0.076)	(0.071)	(0.066)
	150	-0.004	0.003	0.008	-0.008	0.003	0.000
		(0.065)	(0.062)	(0.064)	(0.066)	(0.065)	(0.060)
$s = 12$	50	0.007	-0.000	-0.002	0.002	-0.008	0.009
		(0.060)	(0.061)	(0.065)	(0.060)	(0.066)	(0.056)
	100	0.005	-0.007	0.001	-0.002	0.006	0.002
		(0.062)	(0.063)	(0.058)	(0.060)	(0.057)	(0.052)
	150	-0.004	0.009	-0.009	0.007	0.001	0.001
		(0.055)	(0.063)	(0.061)	(0.057)	(0.054)	(0.050)
$s = 20$	50	0.005	-0.004	0.017	-0.018	0.020	-0.013
		(0.072)	(0.063)	(0.072)	(0.061)	(0.065)	(0.060)
	75	-0.001	0.005	0.007	-0.011	0.002	0.004
		(0.059)	(0.053)	(0.055)	(0.057)	(0.054)	(0.055)
	100	-0.002	-0.001	0.003	-0.005	0.001	-0.001
		(0.055)	(0.060)	(0.059)	(0.058)	(0.056)	(0.058)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 0.90, 0.81, 0.73, 0.66\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n^{1/3}, 3n^{1/3}]$ and 5,000 MC repetitions.

Table C.3-10: Average lead coefficient bias estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j < 0$ for LS method and DGP.2 under systematic skip sampling stock variables.

	n	$\tilde{c}(-1)$	$\tilde{c}(-2)$	$\tilde{c}(-3)$	$\tilde{c}(-4)$	$\tilde{c}(-5)$	$\tilde{c}(-6)$
$s=3$	50	0.000	-0.000	0.001	0.001	0.002	-0.001
		(0.054)	(0.061)	(0.067)	(0.068)	(0.067)	(0.069)
	100	-0.001	0.000	-0.001	0.002	0.001	-0.002
		(0.050)	(0.056)	(0.052)	(0.054)	(0.049)	(0.049)
	150	0.001	-0.001	-0.000	0.000	0.000	-0.001
		(0.049)	(0.054)	(0.056)	(0.058)	(0.056)	(0.058)
$s=4$	50	-0.001	-0.001	0.003	-0.000	-0.000	0.000
		(0.058)	(0.065)	(0.056)	(0.057)	(0.063)	(0.065)
	100	0.001	-0.000	-0.001	0.000	-0.001	0.000
		(0.055)	(0.062)	(0.058)	(0.054)	(0.055)	(0.058)
	150	-0.000	-0.001	0.000	0.000	0.001	-0.001
		(0.051)	(0.052)	(0.050)	(0.051)	(0.052)	(0.051)
$s=12$	50	-0.001	0.002	-0.001	-0.000	0.002	-0.002
		(0.064)	(0.060)	(0.062)	(0.066)	(0.059)	(0.060)
	100	-0.000	0.002	-0.001	0.000	-0.000	-0.001
		(0.060)	(0.059)	(0.054)	(0.059)	(0.050)	(0.062)
	150	-0.001	0.000	0.000	-0.001	0.001	-0.000
		(0.052)	(0.053)	(0.053)	(0.049)	(0.058)	(0.054)
$s=20$	50	-0.000	0.001	0.001	0.001	-0.001	-0.000
		(0.056)	(0.056)	(0.060)	(0.064)	(0.062)	(0.067)
	75	-0.000	-0.001	0.001	0.000	0.000	-0.001
		(0.055)	(0.057)	(0.054)	(0.057)	(0.056)	(0.058)
	100	0.000	-0.000	0.001	-0.000	-0.001	0.000
		(0.050)	(0.058)	(0.050)	(0.053)	(0.053)	(0.057)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 0.90, 0.81, 0.73, 0.66\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$ and 5,000 MC repetitions.

Table C.3-11: Average lead coefficient bias estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j < 0$ for MF-HI method and DGP.2 under simulated GDP sampling scheme. .

	n	$\tilde{c}(-1)$	$\tilde{c}(-2)$	$\tilde{c}(-3)$	$\tilde{c}(-4)$	$\tilde{c}(-5)$	$\tilde{c}(-6)$
$s = 3$	50	-0.079	0.037	0.050	-0.023	0.027	-0.050
		(0.074)	(0.066)	(0.064)	(0.060)	(0.050)	(0.044)
	100	-0.017	0.050	-0.009	-0.029	0.040	-0.041
		(0.076)	(0.072)	(0.071)	(0.070)	(0.065)	(0.056)
	150	-0.023	0.023	0.018	-0.034	0.022	-0.018
		(0.068)	(0.064)	(0.066)	(0.069)	(0.065)	(0.057)
$s = 4$	50	-0.038	0.081	-0.062	-0.045	0.079	-0.029
		(0.069)	(0.071)	(0.069)	(0.064)	(0.054)	(0.045)
	100	-0.058	0.073	-0.055	0.080	-0.085	0.074
		(0.081)	(0.074)	(0.077)	(0.073)	(0.072)	(0.066)
	150	0.028	-0.010	0.006	0.017	-0.015	-0.038
		(0.057)	(0.063)	(0.059)	(0.064)	(0.066)	(0.060)
$s = 12$	50	-0.073	0.180	-0.099	-0.004	0.029	0.081
		(0.063)	(0.062)	(0.063)	(0.059)	(0.060)	(0.053)
	100	-0.037	-0.036	-0.017	-0.042	0.022	0.010
		(0.062)	(0.059)	(0.061)	(0.058)	(0.056)	(0.057)
	150	-0.032	0.039	-0.053	0.053	-0.037	0.024
		(0.057)	(0.061)	(0.054)	(0.058)	(0.061)	(0.058)
$s = 20$	50	0.027	-0.157	0.052	0.063	0.047	0.049
		(0.068)	(0.064)	(0.060)	(0.060)	(0.064)	(0.067)
	75	0.037	-0.034	-0.021	0.080	-0.040	0.011
		(0.058)	(0.058)	(0.060)	(0.052)	(0.060)	(0.058)
	100	0.037	0.014	-0.028	-0.069	0.110	0.010
		(0.062)	(0.056)	(0.057)	(0.063)	(0.064)	(0.057)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 2.90, 5.60, 7.05, 7.34, 6.61\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n^{1/3}, 3n^{1/3}]$ and 5,000 MC repetitions.

Table C.3-12: Average lead coefficient bias estimation and the rejection rate of $\mathbb{H}_0^j : c(j) = 0, j < 0$ for LS method and DGP.2 under simulated GDP sampling scheme.

	n	$\tilde{c}(-1)$	$\tilde{c}(-2)$	$\tilde{c}(-3)$	$\tilde{c}(-4)$	$\tilde{c}(-5)$	$\tilde{c}(-6)$
$s=3$	50	-0.001	0.001	-0.000	-0.000	0.000	0.001
		(0.069)	(0.060)	(0.077)	(0.068)	(0.067)	(0.088)
	100	-0.004	0.002	-0.000	-0.001	0.001	-0.000
		(0.055)	(0.057)	(0.061)	(0.058)	(0.059)	(0.069)
	150	0.003	-0.003	-0.001	0.003	-0.002	0.001
		(0.057)	(0.055)	(0.056)	(0.054)	(0.054)	(0.059)
$s=4$	50	0.002	0.002	-0.004	0.003	-0.001	-0.001
		(0.068)	(0.063)	(0.071)	(0.067)	(0.071)	(0.091)
	100	-0.000	0.001	-0.001	-0.001	0.001	-0.001
		(0.057)	(0.051)	(0.055)	(0.053)	(0.051)	(0.057)
	150	0.001	0.001	-0.001	-0.000	0.001	-0.001
		(0.055)	(0.055)	(0.052)	(0.050)	(0.056)	(0.056)
$s=12$	50	-0.008	0.001	0.006	-0.005	0.001	0.001
		(0.060)	(0.068)	(0.075)	(0.057)	(0.070)	(0.084)
	100	0.002	0.001	-0.002	0.001	0.000	-0.000
		(0.055)	(0.056)	(0.058)	(0.060)	(0.059)	(0.068)
	150	-0.000	-0.001	0.001	-0.001	0.001	-0.000
		(0.053)	(0.056)	(0.056)	(0.052)	(0.052)	(0.056)
$s=20$	50	0.003	-0.002	-0.000	0.001	-0.002	0.001
		(0.065)	(0.064)	(0.076)	(0.068)	(0.070)	(0.083)
	75	0.002	0.001	-0.002	0.000	0.001	-0.000
		(0.057)	(0.056)	(0.062)	(0.053)	(0.055)	(0.069)
	100	0.001	-0.000	-0.001	0.002	-0.002	0.001
		(0.050)	(0.053)	(0.068)	(0.058)	(0.061)	(0.067)

Notes: The subsampling ratio $s = \{3, 4, 12, 20\}$ represent Quarterly - Monthly, Annually - Quarterly, Annually - Monthly, and Monthly - Daily sampling schemes, respectively. We report up to Lag 6, where the true coefficients are equal to $\boldsymbol{c} = \{1.00, 2.90, 5.60, 7.05, 7.34, 6.61\}$. The number of leads/lags was chosen according to the Information Criteria technique presented in Section 2, with $M \in [2n_s^{1/3}, 3n_s^{1/3}]$ and 5,000 MC repetitions.

Bibliography

ADRIAN, T., A. ESTRELLA, AND H. S. SHIN (2010): "Monetary cycles, financial cycles, and the business cycle," Tech. rep.

AKAIKE, H. (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19, 716–723.

AMEMIYA, T. AND W. A. FULLER (1967): "A Comparative Study of Alternative Estimators in a Distributed Lag Model," *Econometrica*, 35, 509–529.

ANDERSON, H. M. AND F. VAHID (2001): "Predicting the probability of a recession with nonlinear autoregressive leading-indicator models," *Macroeconomic Dynamics*, 5, 482–505.

ANDREOU, E., E. GHYSELS, AND A. KOURTELLOS (2010): "Regression models with mixed sampling frequencies," *Journal of Econometrics*, 158, 246–261.

ANDREWS, D. W. K. (1993): "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

ANG, A., M. PIAZZESI, AND M. WEI (2006): "What does the yield curve tell us about GDP growth?" *Journal of Econometrics*, 131, 359 – 403.

ANGELINI, E., J. HENRY, AND M. MARCELLINO (2006): "Interpolation and backdating with a large information set," *Journal of Economic Dynamics and Control*, 30, 2693 – 2724.

BAI, J., E. GHYSELS, AND J. H. WRIGHT (2013a): "State Space Models and MIDAS Regressions," *Econometric Reviews*, 32, 779–813.

——— (2013b): "State Space Models and MIDAS Regressions," *Econometric Reviews*, 32, 779–813.

BAI, J. AND P. PERRON (1998): "Computation and analysis of multiple structural change models," Tech. Rep. 9807, Département de sciences économiques, Université de Montréal, Montreal, Canada.

——— (2003): "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, 18, 1–22.

Bańbura, M. and M. Modugno (2014): "Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data," *Journal of Applied Econometrics*, 29, 133–160.

Barrett, A. B., L. Barnett, and A. K. Seth (2010): "Multivariate Granger causality and generalized variance," *Phys. Rev. E*, 81, 041907.

Beltrão, K. and P. Bloomfield (1987): "Determining the bandwidth of a kernel spectrum estimate," *Journal of Time Series Analysis*, 8, 21–38.

Bernanke, B. S. (1990): "On the predictive power of interest rates and interest rate spreads," *New England Economic Review*, 51–68.

Bernanke, B. S., M. Gertler, and M. Waston (1997): "Systematic Monetary Policy and the Effects of Oil Price Shocks," Tech. rep.

Bonser-Neal, C. and T. R. Morley (1997): "Does the yield spread predict real economic activity? : a multicountry analysis," *Economic Review*, 37–53.

Breitung, J. and B. Candelon (2006): "Testing for short- and long-run causality: A frequency-domain approach," *Journal of Econometrics*, 132, 363 – 378.

Breitung, J. and S. Schreiber (2017): "Assessing causality and delay within a frequency band," *Econometrics and Statistics*.

Bühlmann, P. (1995): "Moving-average representation of autoregressive approximations," *Stochastic Processes and their Applications*, 60, 331 – 342.

——— (1996): "Locally adaptive lag-window spectral estimation," *Journal of Time Series Analysis*, 17, 247–270.

Cai, Z. (2007): "Trending time-varying coefficient time series models with serially correlated errors," *Journal of Econometrics*, 136, 163 – 188.

Cargill, T. F. and R. A. Meyer (1974): "Some Time and Frequency Domain Distributed Lag Estimators: A Comparative Monte Carlo Study," *Econometrica*, 42, 1031–1044.

Chamberlain, G. (1982): "The General Equivalence of Granger and Sims Causality," *Econometrica*, 50, 569–581.

Clements, M. P. and A. B. Galvão (2008): "Macroeconomic Forecasting With Mixed-Frequency Data," *Journal of Business & Economic Statistics*, 26, 546–554.

Cooley, J., P. Lewis, and P. Welch (1969): "The finite Fourier transform," *Audio and Electroacoustics, IEEE Transactions on*, 17, 77–85.

Corbae, D., S. Ouliaris, and P. C. B. Phillips (1997): "Band Spectral Regression with Trending Data," Cowles Foundation Discussion Papers 1163, Cowles Foundation for Research in Economics, Yale University.

——— (2002): "Band Spectral Regression with Trending Data," *Econometrica*, 70, 1067–1109.

Davis, E. P. and G. Fagan (1997): "Are financial spreads useful indicators of future inflation and output growth in EU countries?" *Journal of Applied Econometrics*, 12, 701–714.

Dufour, J.-M., D. Pelletier, and E. Renault (2006): "Short run and long run causality in time series: inference," *Journal of Econometrics*, 132, 337 – 362.

Dufour, J.-M. and E. Renault (1998): "Short Run and Long Run Causality in Time Series: Theory," *Econometrica*, 66, 1099–1125.

Dufour, J.-M. and A. Taamouti (2010): "Short and long run causality measures: Theory and inference," *Journal of Econometrics*, 154, 42–58.

Dufour, J.-M. and D. Tessier (2006): "Short-Run and Long-Run Causality between Monetary Policy Variables and Stock Prices," Working Paper No. 2006-39, Bank of Canada.

Engle, R. F. (1974): "Band Spectrum Regression," *International Economic Review*, 15, 1–11.

——— (1980): "Hypothesis Testing in Spectral Regression; the Lagrange Multiplier Test as a Regression Diagnostic," in *Evaluation of Econometric Models*, National Bureau of Economic Research, Inc, NBER Chapters, 309–321.

Estrella, A. and G. A. Hardouvelis (1991): "The Term Structure as a Predictor of Real Economic Activity," *The Journal of Finance*, 46, 555–576.

Estrella, A. and F. S. Mishkin (1995): "Predicting U.S. Recessions: Financial Variables as Leading Indicators," Working Paper 5379, National Bureau of Economic Research.

Fama, E. F. (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance*, 25, 383–417.

——— (1991): "Efficient Capital Markets: II," *The Journal of Finance*, 46, 1575–1617.

Foroni, C., P. Guérin, and M. Marcellino (2015a): "Markov-switching mixed-frequency VAR models," *International Journal of Forecasting*, 31, 692 – 711.

Foroni, C. and M. Marcellino (2013): "A survey of econometric methods for mixed-frequency data," Working Paper 2013/06, Norges Bank.

Foroni, C., M. Marcellino, and C. Schumacher (2015b): "Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials," *Journal of the Royal Statistical Society Series A*, 178, 57–82.

Galvão, A. B. C. (2006): "Structural break threshold VARs for predicting US recessions using the spread," *Journal of Applied Econometrics*, 21, 463–487.

Geweke, J. (1982): "Measurement of Linear Dependence and Feedback Between Multiple Time Series," *Journal of the American Statistical Association*, 77, 304–313.

Geweke, J. F. (1984): "Measures of Conditional Linear Dependence and Feedback Between Time Series," *Journal of the American Statistical Association*, 79, 907–915.

Ghysels, E. (2016): "Macroeconomics and the reality of mixed frequency data," *Journal of Econometrics*, 193, 294 – 314, the Econometric Analysis of Mixed Frequency Data Sampling.

Ghysels, E., J. B. Hill, and K. Motegi (2016): "Testing for Granger causality with mixed frequency data," *Journal of Econometrics*, 192, 207 – 230.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004a): "The MIDAS Touch: Mixed Data Sampling Regression Models," CIRANO Working Papers 2004s-20, CIRANO.

——— (2004b): "The MIDAS Touch: Mixed Data Sampling Regression Models," CIRANO Working Papers 2004s-20, CIRANO.

Ghysels, E., A. Sinko, and R. Valkanov (2007a): "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 26, 53–90.

——— (2007b): "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 26, 53–90.

Gonçalves, S. and L. Kilian (2007): "Asymptotic and Bootstrap Inference for AR(∞) Processes with Conditional Heteroskedasticity," *Econometric Reviews*, 26, 609–641.

Götz, T. B., A. Hecq, and S. Smeekes (2016): "Testing for Granger causality in large mixed-frequency VARs," *Journal of Econometrics*, 193, 418 – 432, the Econometric Analysis of Mixed Frequency Data Sampling.

Götz, T. B., A. Hecq, and J.-P. Urbain (2014): "Forecasting Mixed-Frequency Time Series with ECM-MIDAS Models," *Journal of Forecasting*, 33, 198–213.

Granger, C. W. J. (1963): "Economic processes involving feedback," *Information and Control*, 6, 28 – 48.

——— (1966a): "The Typical Spectral Shape of an Economic Variable," *Econometrica*, 34, 150–161.

——— (1966b): "The Typical Spectral Shape of an Economic Variable," *Econometrica*, 34, 150–161.

——— (1969): "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, 37, 424–438.

Granger, C. W. J. and P. L. Siklos (1995): "Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence," *Journal of Econometrics*, 66, 357–369.

Guérin, P. and M. Marcellino (2013): "Markov-Switching MIDAS Models," *Journal of Business & Economic Statistics*, 31, 45–56.

Hamilton, J. D. (1994): *Time Series Analysis*, Princeton University Press.

Hamilton, J. D. and D. H. Kim (2002): "A Reexamination of the Predictability of Economic Activity Using the Yield Spread," *Journal of Money, Credit and Banking*, 34, 340–360.

Hannan, E. J. (1963a): *Regression for Time Series in M. Rosenblatt ed., Time Series Analysis*, New York: John Wiley.

——— (1963b): *Regression for Time Series in M. Rosenblatt ed., Time Series Analysis*, New York: John Wiley.

——— (1967): "The Estimation of a Lagged Regression Relation," *Biometrika*, 54, 409–418.

——— (1970): *Multiple Time Series (Wiley Series in Probability and Statistics)*, Wiley.

——— (1973): "Central limit theorems for time series regression," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 26, 157–170.

Harvey, A. C. and R. G. Pierse (1984): "Estimating Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, 79, pp. 125–131.

Harvey, C. R. (1991): "The Term Structure and World Economic Growth," *The Journal of Fixed Income*, 1, 7–19.

Hassler, U. (2011): "Estimation of fractional integration under temporal aggregation," *Journal of Econometrics*, 162, 240 – 247.

——— (2013): "Effect of temporal aggregation on multiple time series in the frequency domain," *Journal of Time Series Analysis*, 34, 562–573.

Hassler, U. and H. Tsai (2013): "Asymptotic Behavior of Temporal Aggregates in the Frequency Domain," *Journal of Time Series Econometrics*, 5, 47–60.

Hidalgo, J. (2000): "Nonparametric Test for Causality with Long-range Dependence," *Econometrica*, 68, 1465–1490.

——— (2005): "A bootstrap causality test for covariance stationary processes," *Journal of Econometrics*, 126, 115 – 143.

Hong, Y., Y. Liu, and S. Wang (2009): "Granger causality in risk and detection of extreme risk spillover between financial markets," *Journal of Econometrics*, 150, 271 – 287, recent Development in Financial Econometrics.

Hosoya, Y. (1991): "The decomposition and measurement of the interdependency between second-order stationary processes," *Probability Theory and Related Fields*, 88, 429–444.

——— (2001): "Elimination of Third-series Effect and Defining Partial Measures of Causality," *Journal of Time Series Analysis*, 22, 537–554.

Hsiao, C. (1982): "Autoregressive modeling and causal ordering of economic variables," *Journal of Economic Dynamics and Control*, 4, 243 – 259.

Jones, M. C. and D. F. Signorini (1997): "A Comparison of Higher-Order Bias Kernel Density Estimators," *Journal of the American Statistical Association*, 92, 1063–1073.

Kuersteiner, G. M. (2010): *Granger-Sims causality*, London: Palgrave Macmillan UK, 119–134.

Ladroue, C., S. Guo, K. Kendrick, and J. Feng (2009): "Beyond Element-Wise Interactions: Identifying Complex Interactions in Biological Processes," *PLOS ONE*, 4, 1–14.

Lemmens, A., C. Croux, and M. G. Dekimpe (2008): "Measuring and testing Granger causality over the spectrum: An application to European production expectation surveys," *International Journal of Forecasting*, 24, 414–431.

Lewis, R. and G. C. Reinsel (1985): "Prediction of multivariate time series by autoregressive model fitting," *Journal of Multivariate Analysis*, 16, 393 – 411.

Lütkepohl, H. (1993): "Testing for Causation Between Two Variables in Higher-Dimensional VAR Models," in *Studies in Applied Econometrics*, ed. by H. Schneeweiß and K. F. Zimmermann, Heidelberg: Physica-Verlag HD, 75–91.

Lütkepohl, H. and D. S. Poskitt (1996): "Testing for Causation Using Infinite Order Vector Autoregressive Processes," *Econometric Theory*, 12, 61–87.

Marcellino, M. (2006): "Chapter 16 Leading Indicators," Elsevier, vol. 1 of *Handbook of Economic Forecasting*, 879 – 960.

Marcellino, M. and C. Schumacher (2010): "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP," *Oxford Bulletin of Economics and Statistics*, 72, 518–550.

Mariano, R. S. and Y. Murasawa (2003): "A new coincident index of business cycles based on monthly and quarterly series," *Journal of Applied Econometrics*, 18, 427–443.

——— (2010): "A Coincident Index, Common Factors, and Monthly Real GDP," *Oxford Bulletin of Economics and Statistics*, 72, 27–46.

McCracken, M. W. and S. Ng (2016): "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business & Economic Statistics*, 34, 574–589.

McCracken, M. W., M. T. Owyang, and T. Sekhposyan (2015): "Real-Time Forecasting with a Large, Mixed Frequency, Bayesian VAR," Working Papers 2015-30, Federal Reserve Bank of St. Louis.

Nishiyama, Y., K. Hitomi, Y. Kawasaki, and K. Jeong (2011): "A consistent nonparametric test for nonlinear causality-Specification in time series regression," *Journal of Econometrics*, 165, 112 – 127.

Phillips, P. C. B. (1987): "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

Pierce, D. A. (1979): "R-squared measures for time series," *Journal of the American Statistical Association*, 74, 901–910.

Politis, D. N. and J. P. Romano (1999): "Multivariate Density Estimation with General Flat-Top Kernels of Infinite Order," *Journal of Multivariate Analysis*, 68, 1 – 25.

Robinson, P. (1983): "Review of various approaches to power spectrum estimation," in *Time Series in the Frequency Domain*, Elsevier, vol. 3 of *Handbook of Statistics*, 343 – 368.

——— (1994): *Time series with strong dependence*, Cambridge University Press, vol. 1 of *Econometric Society Monographs*, 47 – 96.

Robinson, P. M. (1989): *Nonparametric Estimation of Time-Varying Parameters*, Berlin, Heidelberg: Springer Berlin Heidelberg, 253–264.

——— (1991a): "Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models," *Econometrica*, 59, 1329–63.

——— (1991b): "Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models," *Econometrica*, 59, 1329–1363.

Rosenblatt, M. (1984): "Asymptotic Normality, Strong Mixing and Spectral Density Estimates," *Annals of Probability*, 12, 1167–1180.

Schorfheide, F. and D. Song (2015): "Real-Time Forecasting With a Mixed-Frequency VAR," *Journal of Business & Economic Statistics*, 33, 366–380.

Shannon, C. E. (1949): "Communication in the Presence of Noise," *Proceedings of the IRE*, 37, 10–21.

Sims, C. A. (1972): *Are There Exogenous Variables in Short-Run Production Relations*, NBER, 17–36.

——— (1973): "Distributed Lags," Working papers, Center for Economic Research, Department of Economics, University of Minnesota.

Stock, J. H. and M. W. Watson (1989): "New Indexes of Coincident and Leading Economic Indicators," in *NBER Macroeconomics Annual 1989, Volume 4*, National Bureau of Economic Research, Inc, NBER Chapters, 351–409.

——— (1991): *A Probability Model of the Coincident Economic Indicators*, Cambridge University Press, 63–90.

——— (1992): "A Procedure for Predicting Recessions With Leading Indicators: Econometric Issues and Recent Experience," Working Paper 4014, National Bureau of Economic Research.

Stoica, P. and R. L. Moses (2005): *Spectral Analysis of Signals*, Prentice Hall.

Taufemback, C. G. (2018a): "Asymptotic behavior of temporal aggregation in mixed-frequency datasets," Working paper, Universidad Carlos III de Madrid.

——— (2018b): "Nonparametric short- and long-run Granger-causality testing in the frequency domain," Working paper, Universidad Carlos III de Madrid.

Tiao, G. C. (1972): "Asymptotic Behaviour of Temporal Aggregates of Time Series," *Biometrika*, 59, 525–531.

Velasco, C. (2002): "Local Cross-validation for Spectrum Bandwidth Choice," *Journal of Time Series Analysis*, 21, 329–361.

Wahba, G. (1969): "Estimation of the Coefficients in a Multidimensional Distributed Lag Model," *Econometrica*, 37, 398–407.

Wong, W. H. (1983): "On the Consistency of Cross-Validation in Kernel Nonparametric Regression," *Ann. Statist.*, 11, 1136–1141.